A Biophysical Analysis of the Ocr Protein Gel

Richard G. Higham

Doctor of Philosophy

Schools of Physics and Chemistry

University of Edinburgh

2007

Abstract

Ocr is unusual among proteins in its ability to form a transparent gel at high ammonium sulphate concentrations. This transition was investigated using a combination of spectroscopic, microscopic and rheological techniques. It occurs sharply at a concentration of 3.2M ammonium sulphate and is not observed with other types of salt. Rheological measurements showed that rather than precipitating under such conditions, ocr forms a weak viscoelastic gel. Far UV circular dichroism spectra reveal that ocr does not denature in the gel phase, while near UV CD spectra suggest the formation of long, helical structures. Well resolved fibrils were observed using atomic force microscopy. They were over 1μm in length and varied between 2.6nm to 10.4nm in height, corresponding to the thickness and length of the ocr dimer.

Ocr is a highly charged protein (-56e at pH 8) and is shaped like a banana. We argue that it is stabilized in specifically aggregated structures at large salt concentrations by these physical properties. Electrostatic repulsions between proteins are screened by salts, allowing proteins to approach close enough to aggregate. The charge on ocr is high enough to resist such precipitation. However, at 3.2M ammonium sulphate we suggest that the salt molecules bridge neighbouring ocr dimers via hydrogen bonds, connecting amino acid carboxyl groups with the ammonium groups of the salt. The banana-shaped dimers stack on top of each other, forming long helical fibrils that intertwine into a semiflexible network.

Acknowledgements

The three (plus) years of a PhD are a roller coaster ride. The ups and downs you experience, both terrifying and exhilarating, are shared with a great many people. I would like to thank a number of friends and colleagues, without whom this ride would not have been possible.

First and foremost, I would like to thank my supervisors David Dryden and Wilson Poon, for their unwavering support, eternal patience and boundless enthusiasm.

I am very grateful to those who have lent me their time and expertise: Mark Tock for a crash course in molecular biology and protein purification, Lyndsay Sawyer for helping with CD spectroscopy, Cait MacPhee and Nhan Pham for helping with AFM, and Stefan Egelhaaf and the many others who have shared their valuable insights.

Thanks to my office mates past and present (Colin, Otti and Lucio), to the rest of the Rm 4305 'softies' for some quality banter, and to Susie for the many chats and cups of tea! Thanks also to the lab N2.11 crew past and present, Mark, Steve, Mike and Bansi, for keeping me sane when my gels over-ran, my columns dried out, and my bugs mysteriously died!

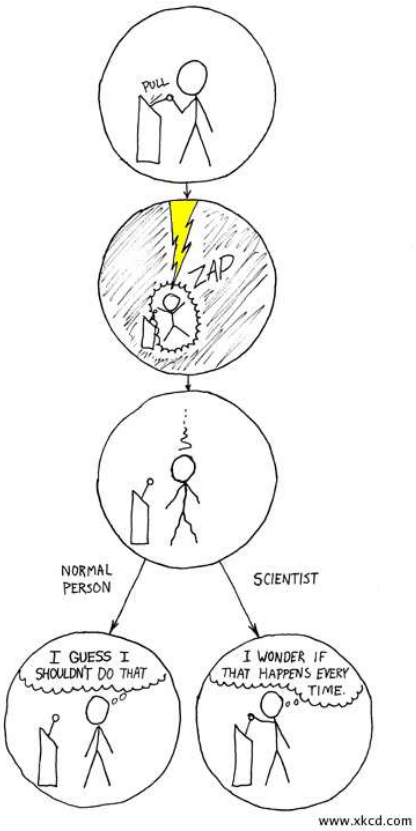
Like any decent fairground ride, upon finishing, you are left with a sublime sense of survival.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Richard G. Higham*)

The Difference



Contents

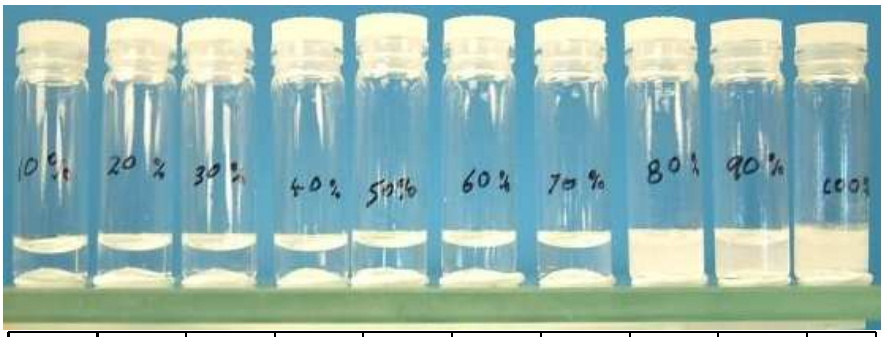
Chapter 1

Introduction

1.1 Bananas, gels, salt and DNA

The subject of this thesis and the link between the above, seemingly unrelated, themes is the protein ocr. While being studied for its important biological role as a DNA structural mimic, unusual physical behaviour in solution was observed. Ocr forms a transparent gel at high salt concentrations where the majority of proteins would normally precipitate. The transition was discovered during a routine precipitation experiment. Ammonium sulphate was added to an ocr solution and the salt concentration was gradually increased. However, ocr did not precipitate, and at 3.2M ammonium sulphate (80% saturation) it rapidly formed a gel. This is in stark contrast with the accepted behaviour of proteins, the majority of which precipitate at concentrations of less than 2M ammonium sulphate. The transition is shown in figure 1.1 and discussed in more detail in chapter 5.

Both the biological function and physical behaviour of ocr stem from the protein's physical properties. Ocr is shaped like a banana and has a large concentration of negative charges on its surface, arranged in a similar manner to the phosphate groups on DNA. These properties allow ocr to mimic a small, bent section of DNA, which is a feature employed by a virus, bacteriophage T7, to help it infect bacterial cells. Viral DNA is usually broken up by a cell's defence mechanisms (restriction enzymes). By appearing physically similar to DNA, ocr disables these defences, preventing the viral DNA from being destroyed (a process called anti-restriction). The high charge density of ocr, vital for its biological function, is also integral to its gel forming capacity.



0.4M	0.8M	1.2M	1.6M	2.0M	2.4M	2.8M	3.2M	3.6M	4M
←—Liquid							Gel —→		

Figure 1.1: Ten solutions of 36μM ocr (1mg/ml) in 20mM Tris-HCL, pH 8, with solid ammonium sulphate added to increasing concentrations. The molar concentration of ammonium sulphate in each solution is shown in the table and the equivalent saturation is written on the bottles. 100% saturation equates to 4M ammonium sulphate. The rapid gel transition occurs at 80% saturation (3.2M). The gels appear opaque due to air trapped in the viscoelastic fluid.

The aim of this thesis is to probe the gel transition of ocr. How does it form? What is the underlying structure? Why is it different from other proteins and protein gels? Such questions were addressed with a number of experimental techniques. The mechanical properties of the salt protein solution were measured to confirm its gel status and investigate its elasticity. This process, called rheology, applies small but precise forces on the gel and measures its response. The results indicate a viscoelastic material (one that can flow while retaining some micro-structure) made from semi-flexible polymers.

Protein structures are highly sensitive to their environment. A spectroscopic technique called circular dichroism (CD) was used to study the structural elements of the ocr proteins (α-helix and β-sheets) under varying salt and pH conditions. Initially it was employed to ensure the salt was not disrupting the native structure of ocr. CD proved unexpectedly sensitive to the gel transition, however, and revealed the rather narrow region of conditions under which a gel forms.

CD also indicated that the structure underlying the gel is chiral in nature. If ocr forms polymers that are twisted into a chiral helix they should be observable on the sub-micron scale. Atomic force microscopy (AFM) has a resolution down to a few nanometers and it was used to build up a detailed topographic image of an ocr gel.

AFM involves gently tracing a sharp tip across the surface of the gel and following all the contours, rather like a stylus following the grooves in a record. The images obtained reveal a network of fibrils formed from the aggregation of ocr proteins into long, stiff polymers.

1.2 Thesis outline

The remainder of this chapter will outline the reasons behind studying this system and introduce the two main concepts of gels and fibrils. Chapter two gives more details about the role of ocr in interactions between viruses and cells, and reviews its physical properties. Chapter three explains the way in which I produced large quantities of ocr from *E.coli* cells by improving existing protocols. In chapter four the three main experimental methods used in this thesis are described. The results of these experiments are reported in chapter five, along with details of the refined process for making ocr gels. Chapter six describes three additional experimental methods and the results obtained using these, which support and complement those detailed in chapters four and five. Chapter seven discusses the results obtained and considers a number of possible models for the ocr gel structure.

1.3 Motivation

Ocr as a protein and as a fibrillar gel demonstrates many interesting features. These are reflected in other natural systems and have potential biological and industrial uses. The main theme of interest in ocr is the study of its anti-restriction properties.

Ocr is currently the largest known DNA mimic and the only anti-restriction protein for which the structure is known. The current research drive is to identify the amino acids essential to its mimicry of DNA. In doing so the structure of its restriction enzyme target (incompletely known so far) can be probed [1]. Researchers hope to use ocr as a framework for building inhibitors of other DNA-binding proteins, with potential diagnostic and medical applications [2].

Nature is replete with examples of self-assembly of biological molecules, from lipids to wood [3, 4]. Some offer beneficial properties like dynamic barriers that can protect sensitive organisms while allowing a controlled supply of required materials

to pass through. Others are involved with human diseases or symptoms like the aggregation of lens proteins causing cataracts. Most are sensitive to their local environment (pH, temperature, ionic strength, solvent composition etc) and are biodegradable, water-based systems. This makes them attractive as potential structures in nanotechnology, for example peptide nanotubules, well-defined crystals or macroscopic membranes.

Amyloid fibrils, in particular, are an extensive field of investigation for protein self-assembly. Given the appropriate conditions most, if not all, proteins can self-assemble into long, stiff fibrils. These are often associated with debilitating conditions like Alzheimer's and Huntington's disease but have a rich and diverse variety of beneficial phenomena in nature. Fibrils occur naturally in cartilage, tendons and skin. Of particular interest is using fibrils as templates for new materials and for controlling assembly of nanoscale molecules [5]. The design of protein fibrils is not as limited as traditional organic polymerisation because the construction of artificial proteins or polypeptides can be precisely controlled from the fundamental amino acid level. Fibrils can be adapted to include non-protein materials such as metals, optically active compounds or other functional groups. They can then be used as a rigid scaffold for nanostructures or as the building blocks for novel biomaterials with specific physical properties [6]. Fibrils can also be precursors to macroscopic systems like gels and liquid crystals.

Gels are an important branch of colloid and macromolecular science as they can form from particles or aggregates of colloidal dimension or from long-chain molecules [7]. They have a theoretical interest in terms of their network topology and the thermodynamics of their formation. There are also a huge range of practical uses for gels. They are a common constituent of lubricants, cosmetics and food products. Fibres formed from the gel state have even been used in the extremely strong material Dynema, used in rock climbing and rope access work [7].

While many proteins form fibrils and gels, they typically involve irreversible aggregation into well-defined structures, or the reversible formation of amorphous aggregates. The reversible formation of fibrils may also occur but at the cost of a structural change in the protein. Ocr self-assembles into fibrils at high salt concentrations without significantly altering its secondary structure. The process is reversible and results in a weak gel, even at low protein volume fraction.

1.4 Anatomy of a gel

One of the most important physical and industrial properties of proteins is their ability to form gels. Heat is often used to induce the transformation but with ocr, salt is used to trigger a gel formation. Protein gels have many interesting properties like their elasticity and water-holding capacity. To adapt these for practical, industrial applications we need to understand how the structure of the gel and its macroscopic properties are related.

Gels occur naturally in biological systems. For example, the membrane of red blood cells is a viscoelastic gel. Its two-dimensional network (the spectrin skeleton) provides a high level of elasticity and stability. Gels are also formed by polysaccharides like cellulose, polypeptides like collagen and denatured proteins like gelatin (unfolded collagen) and β-lactoglobulin [8]. The industrial applications for gels (either protein or synthetic) are many and varied. They are used in cosmetics like toothpaste and shampoo, as thickeners in food or products like paint or washing up liquid, as mobility control agents in oil-field production and they even have a role in drug-delivery systems [9, 10]. Currently the main source of gelation in food products is gelatin, which is bovine in origin.

1.4.1 What makes a gel?

Gels are materials with a very high water content and yet have the structural coherence of a solid, when not being sheared. Gels are normally described in terms of their properties, like being highly elastic and supporting large strains, and defined by results achieved through mechanical testing. For example, from a rheological perspective, the gel phase occurs when the elastic modulus (a measure of the stored energy) is greater than the viscous modulus (a measure of the dissipated energy). Ross-Murphy has extended this definition for polymer gels, requiring a plateau for the elastic modulus over a substantial range of frequencies (figure 5.7) [11]. Rheology is explained in more detail in section 4.1.

A common definition of a gel is a solution of at least two components where one of them has formed a three-dimensional, disordered, flexible structure that spans the system. It exists as an intermediate state between solid and liquid and exhibits mechanical and fluid properties characteristic of each phase [12, 11]. Such networks may

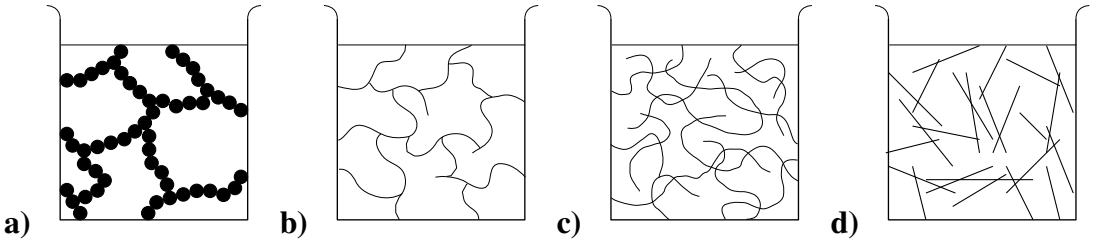


Figure 1.2: Schematics of four possible types of gel: **a)** colloidal, **b)** polymer cross-links, **c)** polymer entanglement network, **d)** rigid rods or fibrils.

be transient or permanent and can form in myriad different ways. Short-range inter-particle attractions between colloids can result in fractal structures or networks (like those shown in figure 1.2a) that confer elasticity to the material, even at low particle volume fractions [13]. At higher densities the colloids may aggregate into a larger, space-spanning structure which restricts the motion of the fluid. Solutions of polymers can also gel by chemically cross-linking their chains (as in figure 1.2b). This is a more permanent effect than simply becoming entangled (as in 1.2c). Here the opposition to flow comes from the frictional resistance of the polymers to moving over each other, as if stirring a bowl of spaghetti [8]. Alternatively, gelation may occur in solutions of rigid rods or protein fibrils when the lengths of the fibers become comparable to their mean separation (figure 1.2d). Gels can also be opaque or transparent, depending on the size and arrangement of their constituents. If globular proteins randomly aggregate the gel formed is generally opaque. Linear aggregation, on the other hand, results in a more transparent gel.

There are too many parameters and situations involved in the gelling process (pH, temperature, solvent character, solute charge and potential binding locations to name but a few) for an all-encompassing model. A number of criteria (proposed by Hermans and refined by Clark, Ross-Murphy *et al.* [14]) are commonly used:

- the growth of a network is disordered and facilitated by units connecting at random
- the concentration of the units controls the degree of cross-linking
- potential cross-linking sites are treated independently
- an equilibrium is set up between creation and dissociation of cross-links

- gel formation is dependent on a critical concentration of cross-links to allow the network to span the sample (a percolation threshold)

These definitions work well for molecular networks that form by a nucleation and growth process, such as the heat gelation of globular proteins or interactions involving reversible associations between polymers, since a number of binding sites are available, through which the subunits may freely interact.

The ambivalent nature of gels comes from the manner in which they are formed. Broadly speaking there are four categories in which gels can fall. They can have particles or polymers as sub-units and can have physical or chemical cross-links. To form a network, discrete particles in a dispersion randomly aggregate into small clusters. Further cluster-cluster aggregation leads to gelation. The strength of connection between colloids will determine the flexibility of the structure [15] and its resistance to deformations. If the structure can accommodate a force without falling apart, as in the spreading of cream cheese, it is said to yield (see section 4.1). If, on the other hand, the structure collapses over length scales larger than the structural units, as in the tearing of Edam cheese, it is said to have fractured [16]. This may occur by the breaking of covalent bonds or the loosening of physical cross-links.

A chemical gel may be formed by polymerization. For example the diallyl bicarbonate monomer can join to other monomers at various points along the molecule [7]. As it grows it forms an interlinked series of polymers. If polymers already exist a network may be formed by joining them with covalent bonds, which result in very stable junction points. This can occur via functional groups along the backbone or using small molecules to form bridges between adjacent chains. For example, polyacrylamide gels consist of covalent bonds between the acrylamide and the crosslinker (*N'*-methylene-bis-acrylamide). The extent of their cross-linking controls the elasticity of the resulting gel. Polyacrylamide gels are used to separate proteins in solution (section 3.1.1) as high molecular weight proteins move at a different rate through the pores of the gel than lower molecular weight proteins. The higher the cross-link density the smaller the pores of the gel, so the gel composition used depends on the sizes of proteins being analysed.

Physical gels form non-covalent bonds through interactions like van der Waals forces, hydrogen bonding or electrostatic interactions, or through phase separations [10]. These are more transient than covalent bonds and the links can be removed by

the same physical means. An example of a substance forming phase-separated micro domains is polyvinyl chloride (PVC) where a number of chains aggregate and give rise to bundle-like, fibrous macromolecules. During a phase transition, crystal particles form and act as junctions, joining different polymer chains. A small section of the chain will be involved with the crystallisation while the remainder forms the network elements of a gel. Electrostatic associations are also a common method of gelation. Polyelectrolytes naturally disperse in solution due to the mutual repulsion of like ions along the chain. By adding salt to the solution the charges become screened and the repulsion decreases. As the chains can move closer to each other ionic bridges are formed between them, eventually leading to a gel. For example, polystyrene sulfonate forms a network with Al^{3+} counter ions [10]. Below a critical polymer concentration the sample precipitates upon addition of Al salt but above this concentration a gel forms as the cross-linked polymers span the sample.

This latter example highlights the important point that network formation is commonly initiated by a dispersion being destabilised. There are many ways to achieve this, including changes in salt, pH, temperature, pressure or the addition of polymers. The most common method of forming gels using proteins like whey and egg lysozyme is through heat [17]. This causes them to denature and, in turn, aggregate. If the protein concentration is high enough the aggregates may span the sample volume, leading to a gel. A similar effect is observed with casein micelles in skimmed milk where van der Waals attraction becomes dominant when the surface molecules that provide steric stabilisation are removed by lowering the pH or by the addition of the enzyme chymosin during cheese production [18].

An example where the addition of polymers induces gelation is seen with poly (methyl methacrylate) polymers and polystyrene particles dispersed in hydrocarbon solvents. The polymers create a depletion potential between nearby particles: when the separation of two particles is less than the size of the polymers they are excluded from that region and the difference in local pressure between the depleteion region and the bulk fluid gives rise to an attractive force between the particles. They aggregate and gradually build up clusters. This aggregation is limited by the diffusion rate of each cluster, which decreases as its volume fraction increases. Eventually the clusters span the sample space and a gel is formed. The depletion attraction between particles are not permanent and clusters will break up with increasing thermal energy. This

competition between aggregation and dispersion is most pronounced at low polymer concentrations where thermal rearrangement dominates and the clusters cannot grow quickly enough to allow gelation [19].

1.5 Protein fibrils

Protein aggregation is often driven by interactions between hydrophobic residues exposed in partially or fully unfolded states. In particular, off-pathway aggregation, where the protein does not fold into its native state, occurs after irreversible thermal denaturation. The protein structure may also be disturbed by increased salt concentrations or changes in pH. Aggregation, though, is not always a random process driven by unfavourable conditions. Certain proteins are able to self-assemble into specific structures, the most common of which are cross-β fibrils. Amyloid fibrils formed from proteins almost exclusively involve β-sheets and an irreversible process. This is in stark contrast to ocr fibrils where there is next to no β-sheet structure and the process is reversible. A number of systems that form fibrils are described below.

1.5.1 Amyloid fibrils

Amyloid fibrils are widely studied as they are involved with a large number of human diseases. For example β-amyloid protein fibrils are a major constituent of insoluble deposits found in the brain of patients with Alzheimer's disease. Other prominent conditions include Huntington's disease (huntingtin protein), Creutzfeldt-Jakob disease (prion protein), senile-systemic amyloidosis (transthyretin protein) and late-onset diabetes involving insulin and amylin fibrils [20, 21]. On a more positive note, fibrils have useful mechanical properties and it has been suggested that they could be used as a functional structure in nanotechnology (for example making conducting circuits from biomacromolecular fibres) or as a bioscaffold in tissue engineering (for example, to control the shape and alignment of cells) [22, 23, 24].

Amyloid proteins aggregate into long, insoluble fibrils. There are no accepted sequence nor structural similarities between proteins that form amyloid fibres; each has unique and well-defined native folds. However, amyloid fibrils have a number of common properties [22, 25, 26]. In particular they all display a characteristic, organised structure consisting predominantly of β-sheets. The fibrils are straight and unbranched

with diameters between 7-12nm and lengths up to several micrometers. They are commonly composed of a number of smaller protofilaments, 2-3nm in diameter. The arrangement of protofilaments varies, sometimes within the same protein sample, to produce structures ranging from flat ribbons to the more commonly observed twisted ropes.

There is also a lag phase (ranging from minutes to days) before well-defined amyloid fibrils are developed, in which aggregates of proteins have been observed by electron microscopy. A nucleation step, similar to that observed in crystal growth, has been proposed for a number of amyloid forming proteins. NMR experiments on peptides indicate the ordered formation of clusters or protofibrils during the lag phase. The rate of fibril formation is seen to be substantially increased by adding pre-formed aggregates of protein [27, 28].

The most common description of fibre secondary structure is the cross-β structure: a polypeptide chain that folds back on itself in a compact s-shape fashion so that the chain is predominantly perpendicular to the fibril axis while the β-sheet hydrogen bonds are parallel to the axis. An alternative to this is the parallel β-helix where β-sheets are arranged on a continuous loop around a cylinder and connected by hydrogen bonds. The interior of the cylinder is filled with hydrophobic side-chains. However, these are only two possible models and there appears to be no single, generic amyloid structure. Neither is β-sheet rich fibril formation limited to a few specific, pathological proteins. Fibrils indistinguishable from amyloids, and unrelated to misfolding protein diseases, have been observed using globular proteins (and even unstructured peptide molecules) with diverse sequences. It appears that aggregation into cross-β sheet fibrils is an intrinsic ability of polypeptide chains, rather than a characteristic of sequence or tertiary fold [20, 29, 21]. The potential for all proteins to form fibrils lies in their shared feature of a peptide backbone, which is central to the intermolecular bonds that stabilize amyloid materials.

A mechanism for amyloid fibril formation, initially proposed for lysozyme mutants by Booth *et al.* [20], is shown in figure 1.3. When the proteins are partially unfolded by heating, unstable β-domains are exposed. These self-associate to produce a stable hydrogen-bonded β-sheet structure, which may act as a template or seed, encouraging the addition of more polypeptide chains to the growing structure. This process will eventually lead to the formation of long, insoluble amyloid fibrils, like those shown

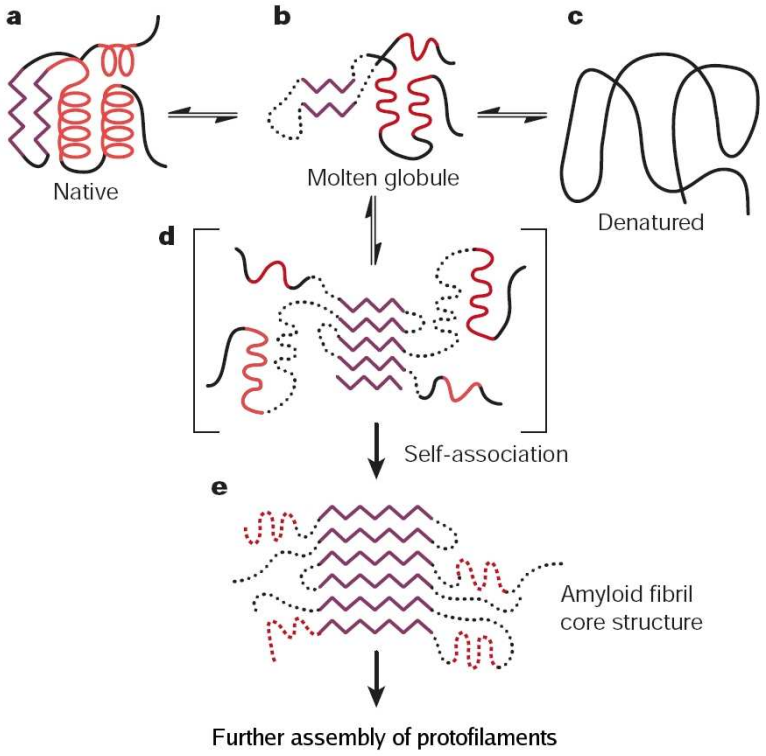


Figure 1.3: A mechanism proposed by Booth *et al.* [20] for lysozyme amyloid fibril formation. A partially unfolded protein (b), with a tertiary structure between that of the native (a) and fully denatured (c) states, self-assembles with other proteins in similar states through their β-domains. As more proteins associate with the oligomer a stable, mainly β-sheet structure develops that forms the core of amyloid fibrils (e). In the diagram, reproduced from [20], purple represents β-sheet structures, red, helical structures and dotted lines, undefined structures not involved in the cross-β fibril core.

in figure 1.4. However, this model does not explain the formation of amyloid fibrils from proteins without any native β-sheet domains. For example, myoglobin has only well-defined α-helices in its secondary structure, yet is able to form fibrils with core structures indistinguishable from that of cross-β fibrils [30]. At high temperature and pH, conditions that destabilise the native secondary structure, proteins like myoglobin and insulin are able to refold its polypeptide chain into β-domains that associate into fibrils.

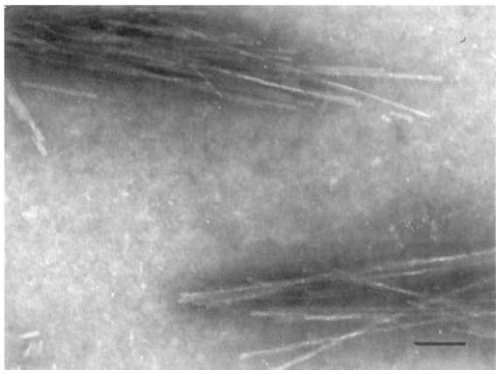


Figure 1.4: Electron micrograph of fibrils formed from 1mg/ml Ile56Thr lysozyme. The protein was incubated in 10mM HEPES, 1M LiCl, pH 8.0 at 4°C for 14 days. Scale bar: 100nm. Reproduced from [20].

1.5.2 Polyglutamine

In Huntington's disease (along with other similar conditions) pathological huntingtin proteins have an expanded glutamine amino acid repeat in their sequence. This causes the protein to aggregate into insoluble filaments and fibrils.

Proteins with a large number of contiguous glutamine residues (generally more than 37) are able to form 3nm wide filaments (similar to amyloid protofibrils) [31]. The polymerisation is self-initiating, concentration and repeat-length dependent and involves a nucleation step. Experiments have been conducted on peptides like GST-HDex1-Q51 which has 51 glutamine residues [32]. They are incubated at 37°C for 24 hours in a suitable buffer (for example, 10mM Tris-TFA, pH 7.0) and after a lag phase due to nucleation there is a rapid aggregation involving almost all the monomers to form filaments. These short filaments may form a mesh or network structure. More commonly, however, they nucleate into longer threads that may exist in isolation or align to become part of a broad ribbon around 40nm wide and from 100nm to several micrometers long. Only rarely do they aggregate into the trademark twisted motif of amyloid fibrils. Despite this they share other amyloid hallmarks, like an extensive β-sheet structure.

CD spectra obtained for polyglutamine are similar to those of amyloid fibrils. A large change in the signal at far UV wavelengths showed a change from random coil to β-sheet motifs. The exact arrangement of β-sheets in polyglutamine fibrils is unknown. The two most commonly described models are a cross-β structure or parallel β-helix

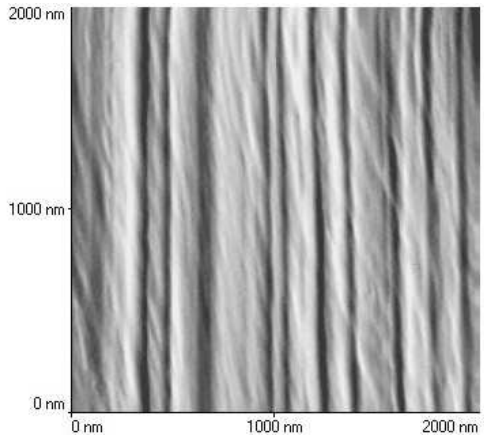


Figure 1.5: AFM image of a Bombyx mori silk fiber's fibril structure. Fibrils are around 100nm in diameter and form a bundle around 10μm wide. Reproduced from [35]

[33]. In the former the sheets either arrange in antiparallel layers or a single sheet folds back on itself in a compact, snake-like manner. In the latter the β-sheets loop continuously round in a cylindrical fashion with the interior filled with side chains and the ladder of β-sheets connected by hydrogen bonds.

1.5.3 Silk fibrils

Silk proteins have an amino acid structure that predominantly consists of repeated regions of alanine, sandwiched between regions of glycine [34]. When spiders produce threads of silk for webs the proteins transform irreversibly from individual, soluble proteins to micrometer long, insoluble fibers, around 10-20μm in diameter. Scanning and transition electron microscopy revealed that fibers are made up of bundles of smaller fibrils around 100-150nm wide [35, 36]. They are long, straight and aligned parallel to the fiber axis. X-ray scattering and NMR (nuclear magnetic resonance) experiments have shown the fiber structure to be mainly β-sheets orientated parallel to the axis of the fiber. The polyalanine regions are largely responsible for the β-sheets while the intermediate glycine rich regions provide elasticity to the fiber. Figure 1.5 shows a portion of a silk fiber with the fibrils clearly visible. Some fibrils display helical features as a result of their underlying β-sheet structure.

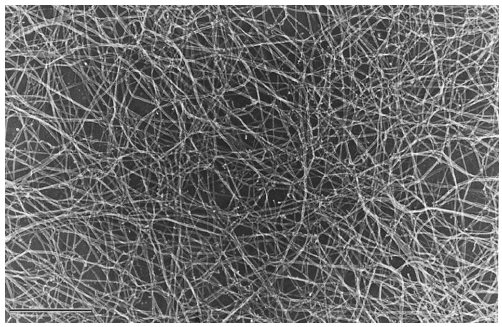


Figure 1.6: Electron micrograph of 1mg/ml actin polymerized by 2mM $MgCl_2$. Filaments 10nm in diameter with a right-handed helical periodicity of 37nm are produced and form an entangled network of semiflexible actin polymers. Scale bar 1μm. Reproduced from [43].

1.5.4 Actin filament gels

Actin is a highly abundant protein found in most living organisms, in particular eukaryotic cells where it comprises between 5-10% of the total protein mass. It self-assembles into structures that are involved in cell locomotion, motility of molecules within the cell (like myosin), cytokinesis (division of cell into two daughter cells) and the cytoskeleton, where it is a major structural component [37, 38]. In particular it is largely responsible for the viscoelastic behaviour of cells under external mechanical force.

Monomeric, globular actin consists of 375 amino acids, has a molecular weight of 42 kDa and a charge of -10e at pH 7. In the presence of cations found under physiological conditions (Mg^{2+}, Ca^{2+} and K^{+}among others) it polymerizes into right-handed helical filaments around 5-10nm in diameter and several micrometers in length. The process occurs at room temperature over a time scale of minutes to hours. The monomers connect non-covalently through electrostatic and hydrophobic interactions. The filaments get entangled and form an irregular mesh-like network (figure 1.6), which has weak characteristics of a gel [39, 40, 41, 42].

Actin filaments though cannot be solely responsible for maintaining the structural integrity of eukaryotic cells. Cytoplasmic actin gels have elastic moduli in the range 100-1000 Pa and can withstand shear stresses of up to 1000 Pa, compared to an elastic modulus of only 0.1 Pa and a breaking stress of <0.1 Pa for a solution of entangled filaments [41]. In addition to actin there are several proteins in vivo that can bind

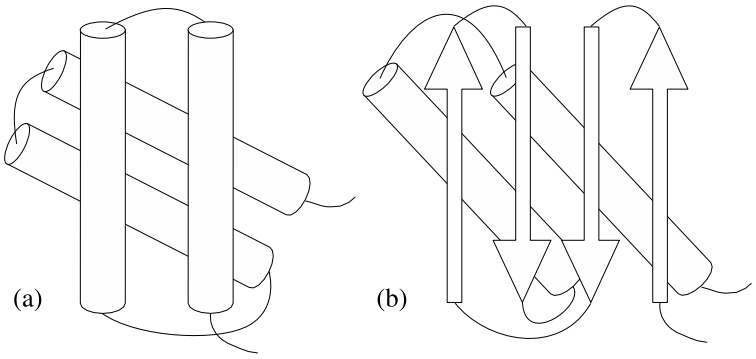


Figure 1.7: The two possible conformations of the prion protein: (a) the benign form consisting of four alpha helices (cylinders) connected by random coil segments; (b) the pathogenic form consisting of two alpha helices and two co-aligned stretches of beta sheet (each denoted by counter-directed arrow pairs). Diagram based on [44].

to actin filaments and form much stronger gels. These actin-binding proteins cross-link and bundle actin filaments into highly viscoelastic networks. In contrast to the bonds between monomers, the bonds formed by actin-binding proteins are rigid and irreversible. They convey a considerable elastic stiffness to the gel structure.

1.5.5 Prion proteins

The primary structure of the prion protein can adopt two distinct secondary and tertiary structure conformations. In its benign form the prion protein exists as four alpha helices but in its pathogenic form the conformation changes to two alpha helices and two co-aligned stretches of β-sheets (figure 1.7) [44]. This motif is known as scrapie and is highly dangerous as the proteins not only readily aggregate into insoluble fibrils but also induce benign forms of prion into adopting the β-sheet conformation. This is the cause of bovine spongiform encephalopathy (BSE), the brain disease in cattle, and scrapie in sheep.

1.6 Banana shaped molecules

Achiral, bent-core molecules can form a liquid crystalline phase with macroscopic chiral properties. Interest in this has been driven by its potential use in electro-optical display devices as chirality is an important feature of liquid crystal displays. The exact

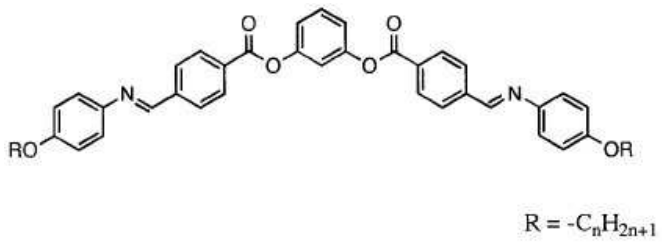


Figure 1.8: The banana-shaped molecule of the 1,3-benzene bis [4-(4-n-alkoxyphenyliminomethyl) benzoate] series (P-*n*-O-PIMB, where n=8, 10, 12 and 16). Reproduced from [46].

manner in which these achiral molecules aggregate to produce chirality in the system is unknown and a number of models have been proposed. The banana-shaped molecule commonly employed in research belongs to the 1,3-benzene bis [4-(4-n-alkoxyphenyliminomethyl) benzoate] series (figure 1.8). When cooled from the high-temperature isotropic phase they spontaneously form fluid smectic phases with chiral layers [45]. They demonstrate only weak birefringence, no anisotropy and a CD peak at $\sim$400nm, indicating the presence of a helical structure [46].

The most common description of these chiral layers involves the molecules aligning in the direction of bending. These arrange into rows that are efficiently packed into a layer. When they are tilted from the smectic plane normal by the same amount, a macroscopic chirality is induced in the domain (figure 1.9) [47, 48, 49, 50]. The tilt arises naturally but can be switched, along with the chirality, by applying an external electric field.

Chirality may also be achieved by the molecules arranging into helical structures or being twisted themselves. Different solvents will affect the shape and polarity of the molecules allowing a range of possible conformations. Goodby *et al.* observed helical ribbons (5-20μm across) forming after cooling from the isotropic phase to the liquid crystal phase, where they arranged into layers [51].

Thisayukta *et al.* suggest alternative models of helicity [52, 46]. The arms of each molecule could be twisted by different angles with respect to the central link. Less intrusive would be for the molecules to align side by side, each rotated about the common axis to form a helix (figure 1.10 b). A chiral structure may also form from the helical rotation of contiguous liquid crystal domains with respect to each other where the helical axis is parallel or perpendicular to the bend direction (figure 1.10 a).

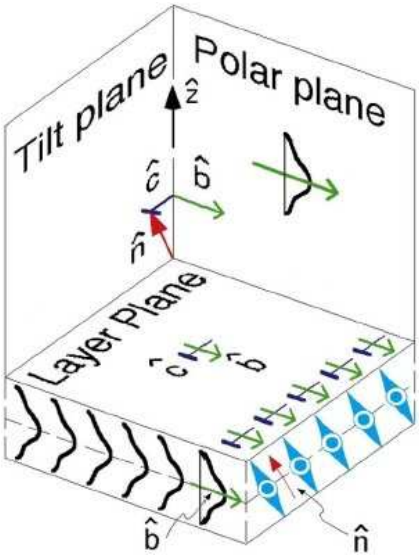


Figure 1.9: Banana-shaped molecules aligning along their bow axis in layers where each is tilted from the plane normal by the same amount. This results in a chiral, smectic liquid crystal domain. Reproduced from [45].

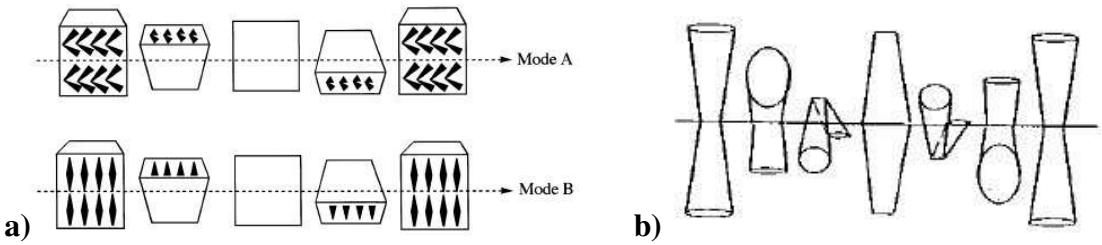


Figure 1.10: Models proposed by Thisayukta *et al.* for chiral structures formed from banana-shaped molecules [52, 46]. Helices are formed from smectic domains **(a)** or the individual molecules themselves **(b)**. In the domains **(a)** the molecules are aligned along their bend direction and rotated about an axis parallel (mode A) or perpendicular (mode B) to this.

1.7 Concluding remarks

These are just a few examples of the large number of systems in nature that exhibit fibrillar properties and form gels. The structure formed by ocr displays similarities with many of them but significant differences remain, suggesting a novel gelation mechanism. Over the course of this thesis we will discover how ocr is able to combine its unusual physical properties in a unique manner.

Chapter 2

Biophysics of ocr

2.1 Proteins

A protein is a collection of amino acids arranged in a specific order along a linear polymer chain, which is bent and twisted into a unique three-dimensional structure. There are 20 naturally occurring amino acids. Each consists of a carboxyl and amide group connected to a central carbon atom (the α-carbon) and are distinguished by a particular side chain (figure 2.1 a) [53, 54]. The side chains confer polar, hydrophobic, acidic or basic properties to the amino acids. Amino acids join together through covalent peptide bonds to form the protein polypeptide backbone.

There are four different levels of structure formed by amino acid chains, which helps to characterise proteins.

- **Primary structure**
  The order in which the amino acids are linked. This is also called the protein sequence and could be several hundred amino acids in length.. The sequence of amino acids is described from the N-terminal end (NH_2 group) to the C-terminal end (COOH group).

- **Secondary structure**
  These structure elements occur locally at various points along the sequence. They involve a number of amino acids forming a regular, recurring motif due to the interactions (mainly hydrogen bonds) between neighbours. The two main forms of secondary structure are alpha-helix and beta-sheets. Alpha helices involve hydrogen bonds between nearby amino acid residues while beta-sheets

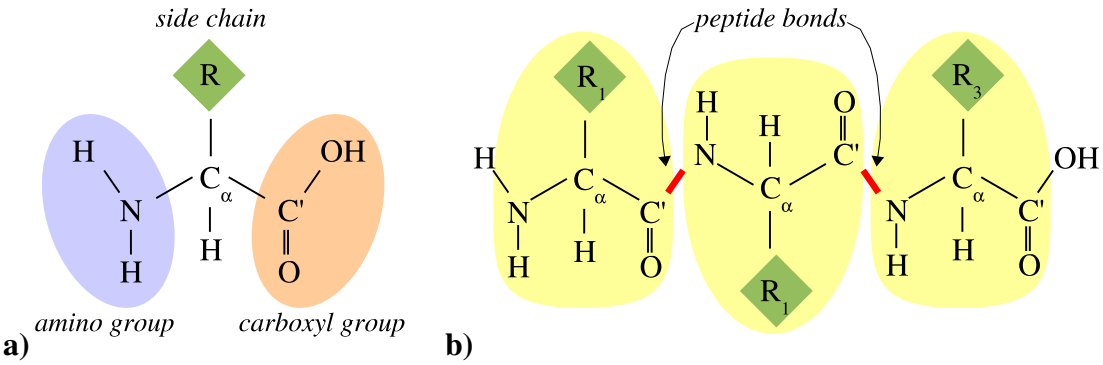


Figure 2.1: **a)** The elements of an amino acid: the central α-carbon, side chain (R), amino and carboxyl groups. **b)** The polypeptide chain that forms the primary structure of a protein.

are formed by hydrogen bonding between sections of the polypeptide chain not necessarily contiguous.

- **Tertiary structure**
  This is the overall three-dimensional shape of the protein, when it has adopted a stable conformation. The polypeptide chain and secondary structures bend and twist into a tightly folded structure. The exact process by which this happens is not known but it involves a number of interactions both within the molecule and between amino acid side chains and surrounding water molecules. For example, hydrogen bonds, hydrophobic interactions between side chains that group together to exclude water, disulphide bonds between cysteine groups, attractive van der Waals forces, or electrostatic interactions between groups like lysine and glutamic acid or between polar side-chains and water dipoles. The naturally occurring form of a protein is called its native tertiary structure.

- **Quaternary structure**
  If a number of polypeptide chains fit together to form a composite molecule this is called its quaternary structure. In this case the individual proteins are called monomers; the composite structure is the multimer.

Protein structure is closely related to function. The practically limitless number of possible combinations of amino acids gives rise to a large array of functions. Proteins called enzymes function as biochemical catalysts. Haemoglobin is used to transport

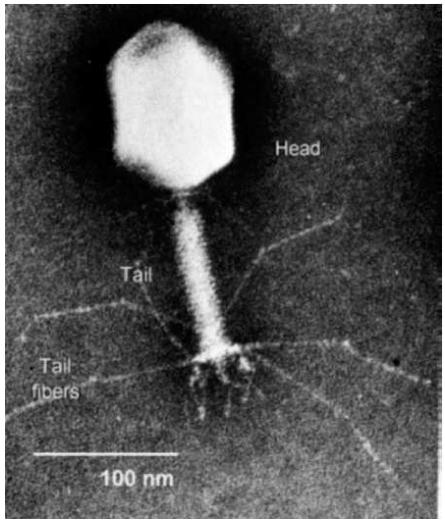


Figure 2.2: An electron micrograph of a T4 bacteriophage from the Universal Virus Database (http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB).

oxygen to tissues while serum albumin can be partially attached to drugs and guide them to the desired location. They can provide mechanical motion (for example actin and myosin in muscle) and structural support (like strengthening skin and bone with collagen). They are also involved in the immune system through antibodies.

2.2 Bacteriophage

Bacteriophage are viruses that infect bacteria in order to replicate, destroying the cell in the process. The name comes from the Greek work *phagein*, to devour. A common group of bacteriophage used in research is the T series[1], which infect *Escherichia coli (E. coli)*. Each phage consists of a protein capsule containing a highly compact, double strand of DNA. The head is attached to a tubular tail, at the end of which may be a number of legs. Figure 2.2 shows an electron micrograph of a T4 bacteriophage. It floats around dormant until sensors on the ends of the legs recognise a host cell. The bacteriophage is triggered into action and it binds to the surface of the larger organism. Using its tail it then injects its DNA into the cell. The natural processes in the cell are subverted and it replicates the viral genome in place of its own. Eventually so many copies of the virus are produced that the cell bursts, releasing the new bacteriophage

[1] T stands for “type”

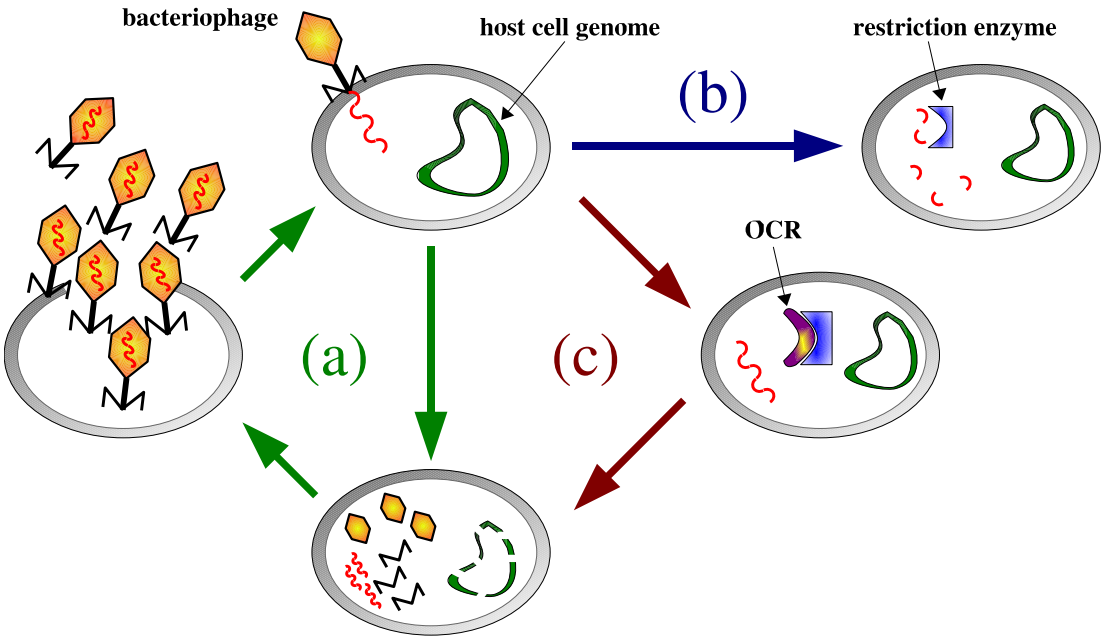


Figure 2.3: a) The virus cycle. A bacteriophage injects its DNA into the host cell. It is replicated and transcribed, then viral mRNAs are translated and copies of the bacteriophage are made. The cell eventually bursts and the bacteriophage are released, free to repeat the process with a new cell. b) The cell defence mechanism: a restriction enzyme chops up viral DNA. c) The bacteriophage counter attack: ocr, an anti-restriction protein. This binds to the restriction enzyme, disabling it and allowing the viral DNA to infect the cell.

into the environment (summarised in figure 2.3a).

2.2.1 Restriction

To defend itself against phage attack, bacteria have developed restriction-modification (R-M) systems. These have a dual purpose: to destroy viral DNA invading the cell (restriction) while protecting the bacteria's own DNA (modification). Host DNA is labelled by an enzyme called a methyltransferase (MTase) by attaching a methyl group to a specific adenine or cytosine on specific sequences within the genome. This distinguishes it from unmethylated, foreign DNA, which is cut, or cleaved, by restriction endonuclease (REase) [55]. These two enzymes are combined into one unit in type I R-M systems. Once a type I R-M enzyme locates a piece of DNA it will function

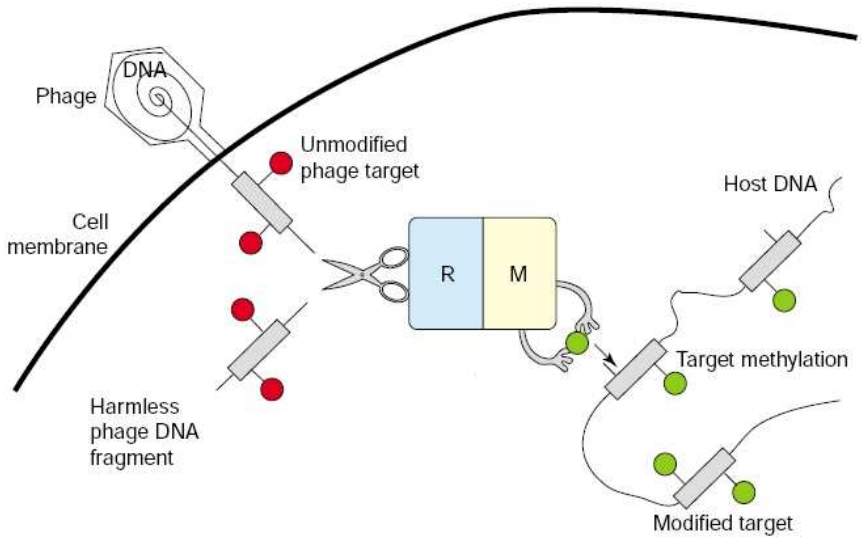


Figure 2.4: Diagram illustrating the function of a type I Restriction-Modification enzyme. These enzymes recognise a specific target sequence on DNA (shown as a grey rectangle) and their consequent behaviour depends on its state of methylation. Newly replicated bacterial DNA is hemimethylated (shown as a single green circle on host DNA target sequences) and the enzyme modifies the sequence by a performing a second methylation. Foreign DNA, invading the cell, is not methylated at all (shown as red circles on phase DNA target sequences)and is therefore cleaved by the enzyme into harmless fragments, preventing the phase from infecting the cell. Diagram reproduced from [1].

as either an MTase or REase depending on the methylation status of the DNA (figure 2.4). Following replication, host DNA only has one methyl group on the sequence recognised by R-M enzymes and is termed hemimethylated. In this case R-M enzymes will add another methyl group making the DNA fully methylated, while unmethylated (foreign) DNA is cut into random fragments. The type I R-M enzyme remains bound to the target site where it initially attached while pulling in (translocating) the foreign DNA from both sides simultaneously. When it collides with another bound R-M enzyme the double strand of DNA is cut at the collision point [56]. Sometimes phage DNA escapes restriction and is accidentally labelled with a protective methyl group, along with the host chromosome [57, 58]. Although they are not 100% effective, R-M enzymes like EcoKI can limit (“restrict”) phage propagation by up to a factor of 10^8.

2.2.2 Anti-restriction

In response bacteriophage have evolved a number of ways to evade destruction by R-M systems. Some try to avoid a confrontation by changing their DNA sequence so R-M enzymes cannot detect them or by having proteins that enter the cell along with the DNA and attaching to the places which are normally targeted by R-M systems. Others prefer to tackle the enzymes straight on by subverting them into methylating foreign DNA or by inhibiting them altogether [1, 59]. Of particular interest are the bacteriophages T3 and T7, which encode the anti-restriction protein ocr to protect their own DNA.

2.3 Ocr

The ocr protein[2] plays a crucial role in the interaction between cells and bacteriophage. It is the first protein expressed by bacteriophage T7 following the infection of *E. coli* [60]. Once produced, each ocr protein binds to an R-M enzyme, repressing its restriction and modification abilities and preventing T7 DNA from being destroyed by the R-M system [57, 56]. The viral DNA can then be freely transcribed. The inhibition of R-M enzymes is enabled by the physical similarity of ocr to B-DNA (figure 2.8). Ocr has been shown to mimic structural and electrostatic properties of DNA. This allows it to bind tightly to the R-M enzymes blocking their binding to viral DNA and rendering them inert [61, 62]. The binding affinity between ocr and the type I enzymes EcoKI and EcoBI is so favourable (50-times greater than for type I/DNA binding [1]) that ocr competes with DNA for the binding site on the enzymes and even causes them to dissociate from DNA [63, 55].

2.3.1 Primary structure

There are 116 amino acids in an ocr monomer. They are ordered as shown in figure 2.5 and the amino acids present are summarised in table 2.1. Only one of the 20 natural amino acids is not present in ocr: cysteine. This means no disulphide bridges can be formed. There are a large number of aspartic and glutamic acids in the polypeptide chain. These give ocr its large negative charge (see section 2.6). Each monomer also

[2] ocr stands for Overcome Classical Restriction.

```
AMSNMTYNNV FDHAYEMLKE NIRYDDIRDT DDLHDAIHMA ADNAVPHYYA
DIFSVMASEG IDLEFEDSGL MPDTKDVIRI LQARIYEQLT IDLWEDAEDL
LNEYLEEVEE YEEDEE
```

Figure 2.5: ocr's 116 amino acid sequence (primary structure). The amino acid that each letter represents can be found in table 2.1. Sequence from Dunn *et al.* [65].

48 Hydrophobic			**29 Polar**			**40 Charged**		
A	Alanine	10	N	Asparagine	6	D	Aspartic Acid $^{(-)}$	17
G	Glycine	2	Q	Glutamine	2	E	Glutamic Acid $^{(-)}$	17
I	Isoleucine	9	H	Histidine	4	R	Arginine $^{(+)}$	4
L	Leucine	10	S	Serine	4	K	Lysine $^{(+)}$	2
M	Methionine	6	T	Threonine	4			
F	Phenylalanine	3	W	Tryptophan	1			
P	Proline	2	Y	Tyrosine	8			
V	Valine	5	C	Cysteine	0			

Table 2.1: The number of each amino acid present in an ocr monomer. 19 of the 20 natural amino acids are present, the exception being cysteine. Note the high content of aspartic and glutamic acid that gives ocr its characteristic charge.

has 1 tryptophan (W) and 8 tyrosine (Y) amino acids, which contribute to its extinction coefficient of $15547 M^{-1}\text{cm}^{-1}$ at 280nm (section 3.1.2). Ocr has a molecular weight of 13678 Daltons[3]. The properties quoted above are for an ocr monomer and the quantities should be doubled for a dimer. By using a combination of centrifugation and gel-filtration (see section 3.3.4.1) Mark and Studier showed that ocr exists in aqueous solutions predominantly (>99%) as a dimer, both at high ionic strength and under physiological conditions [60]. This was confirmed by Blackstock *et al.* using light scattering [64].

2.3.2 Crystal structure

The polypeptide chain for ocr folds into four alpha helices per monomer (with no beta sheets). Two monomers join to form the structure shown in figure 2.7. At pH 8.0 the ocr dimer has 56 negative charges and resembles a 24 base-pair piece of B-DNA,

[3]Daltons are equivalent to atomic mass units (amu).

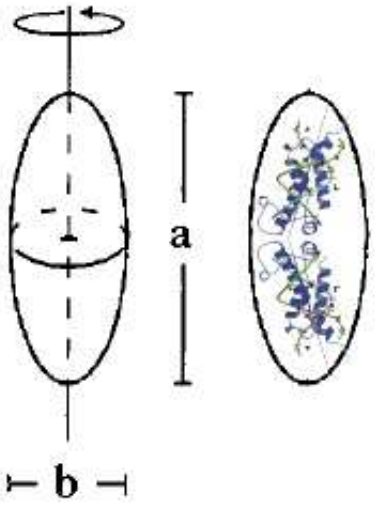


Figure 2.6: Sketch of a prolate ellipsoid (from [64]) and one with the ocr dimer included. For ocr the dimensions are a=10.4nm and b=2.6nm.

bent in the middle at 33° (figure 2.8). The negative carboxyl groups are arranged on the surface of ocr so they mimic the distribution of phosphate groups on B-DNA and complement the target recognition domains of the type I RM enzymes [55]. The size and shape of ocr are also consistent with a bent DNA molecule bound by a restriction enzyme. Dynamic light scattering reveals that a prolate ellipsoid with dimensions of 10.4nm by 2.6nm provides the best description for the shape of the ocr dimer (figure 2.6) [64]. EcoKI binds to a section of double helix DNA 9nm long and 2nm wide, causing it to bend by a similar amount as the bend in the ocr molecule [66].

2.4 Introducing pH and protein charge

The charge carried by a protein depends on the pH of its solvent. The pH is a measure of the concentration of hydrogen ions in a solution and can be defined as:

$$pH = -log_{10}[H^+] \tag{2.1}$$

where $[H^+]$ represents the molar concentration of hydrogen ions. These bind to and dissociate from the acidic amino acid side chains and the amino and carboxyl termini of the protein chain, controlling their ionization. To maintain a constant charge on a protein the pH of the solution must also remain constant, an effect called buffering.

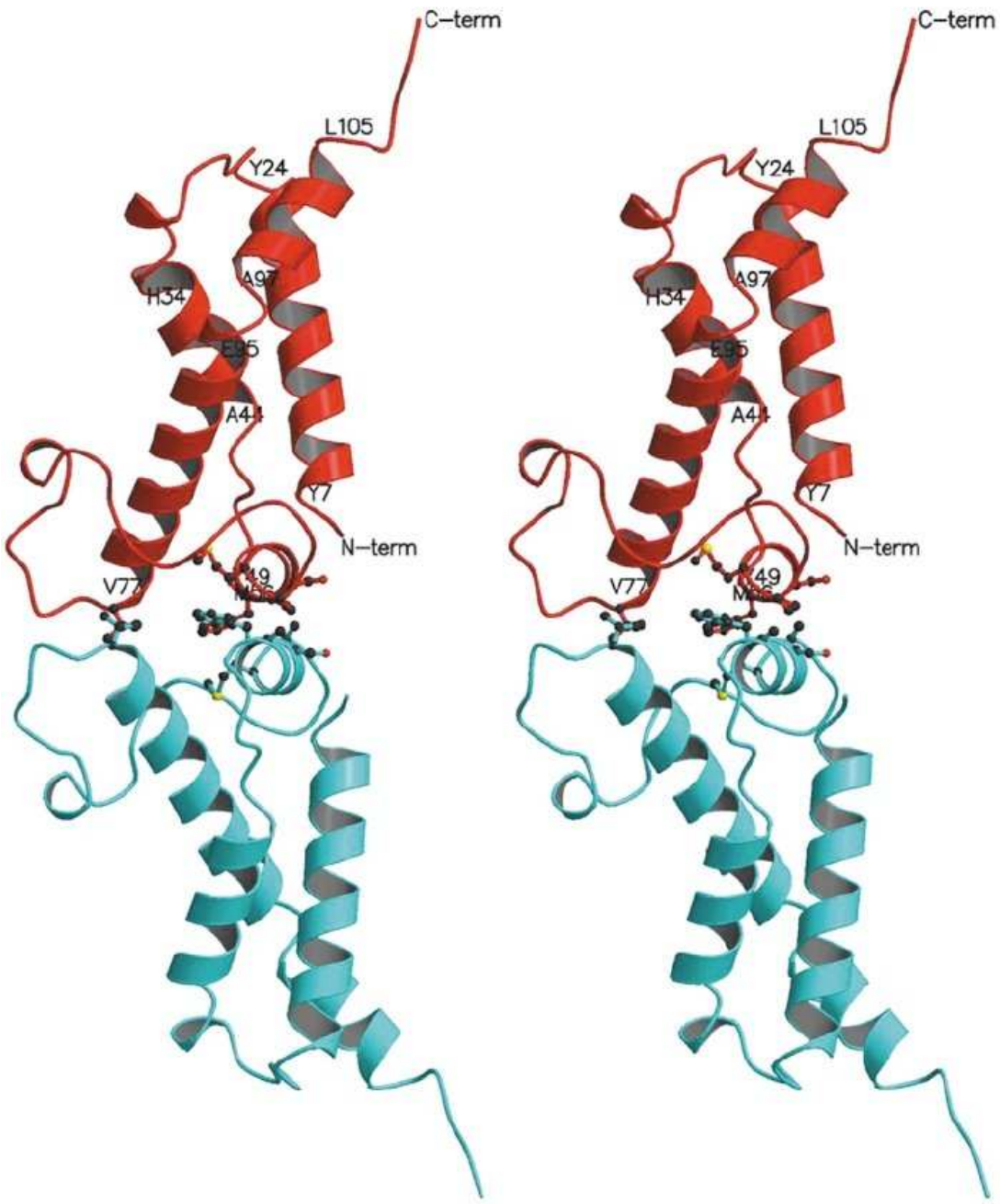


Figure 2.7: A stereo image of the ocr dimer (cross your eyes to view a 3D rendering of the molecule). The two monomers are highlighted in red and blue. The amino acids marking the ends of each alpha helix are labelled: Y7 to Y24, H34 to A44, Y49 to M56 and V77 to L105 [66].

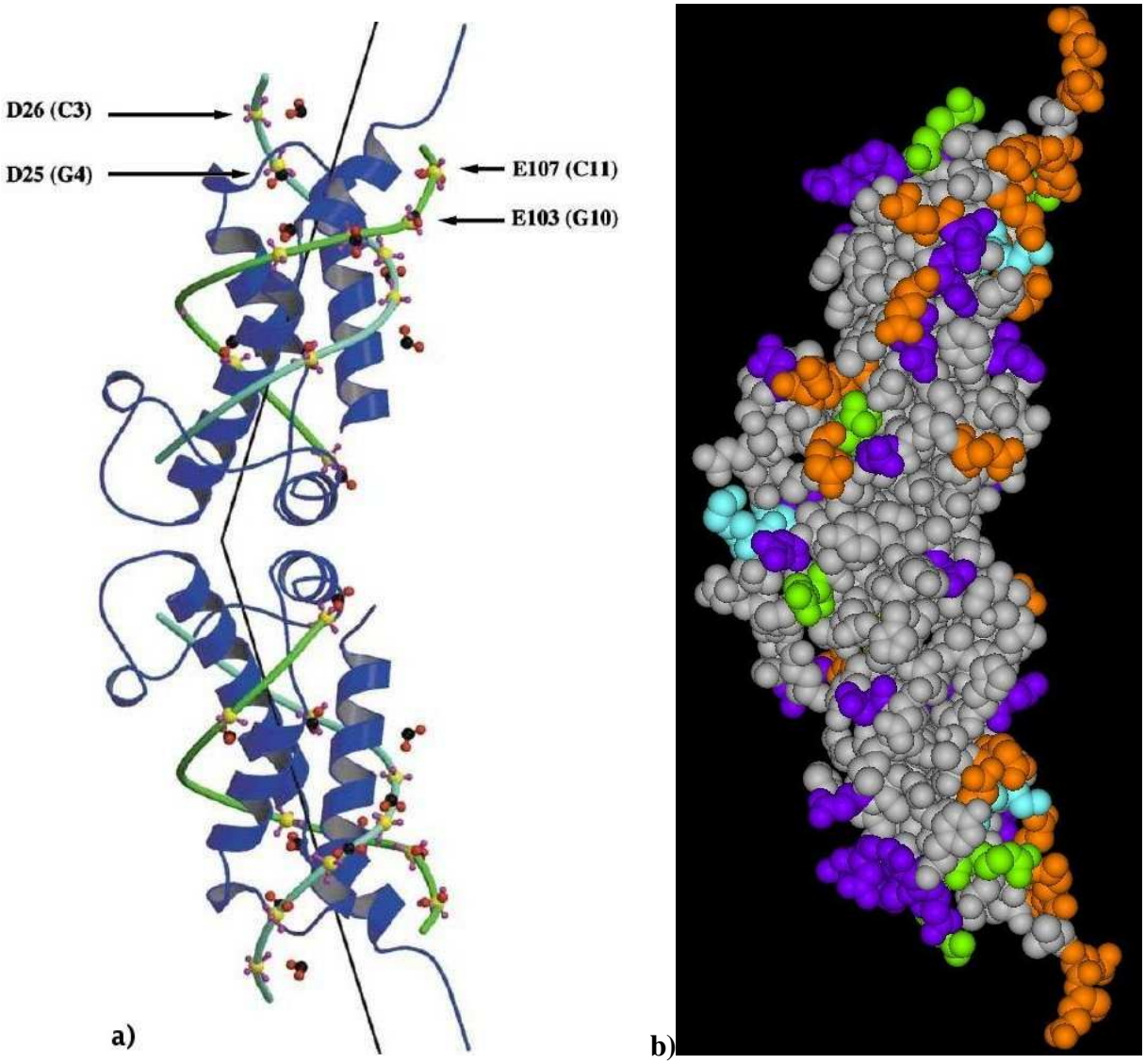


Figure 2.8: **a)** The ocr dimer (dark blue) with a bent double helix (green) superimposed to highlight the similar location of charged groups on each molecule. The DNA phosphate groups are coloured yellow and purple (C=cytosine, G=guanine). The ocr carboxyl groups are coloured red and black (D=aspartic acid, E=glutamic acid). The intersection between the top and bottom DNA helices gives a bend angle of 33.6°. Reproduced from [66]. **b)** The ocr dimer showing the charged amino acids: purple = aspartic acid (-), orange = glutamic acid (-), blue = lysine (+) and green = arginine (+). Molecule plotted with Accelrys ViewerLite 5.0 (http://www.accelrys.com) .

See figure 2.18 for acidic, basic, hydrophobic and hydrophilic groups on ocr.

At a certain pH value there are an equal number of positive and negative charges on the protein. The protein is at its isoelectric point, pI, where the net charge is zero. For pH values below its pI the protein is positively charged, while at values above its pI it is negatively charged. Proteins are normally buffered within the pH range 6-8. This is close to the physiological pH of most animal cells (7 - 7.5) and restricts denaturation. A plot of the charge on a protein against solution pH is called a titration curve. These will be discussed in section 2.6 after an introduction to the concept of screening in section 2.5.

2.4.1 Acids and bases

The common definition of acids and bases follows the theory of Brønsted-Lowry[4] [67]. An acid is a chemical compound that may donate protons (or equivalently hydrogen ions, H^+) to another compound called a base, which accepts protons. This can be summarised as follows:

$$HX + Y^- \rightleftharpoons HY + X^-$$

where HX and HY are Brønsted-Lowry acids, and X^- and Y^- are Brønsted-Lowry bases. An equilibrium occurs when an acid (HA) is dissolved in water, which accepts a proton from the acid and therefore acts as a base. This is expressed in equation 2.2. In the reverse reaction the ion A^- accepts a proton to return to its HA form and acts as the base. A^- is said to be the conjugate base of the acid HA and H_3O^+ the conjugate acid of the base H_2O.

$$HA + H_2O \rightleftharpoons H_3O^+ + A^- \tag{2.2}$$

The equilibrium constant for this reaction is the acidity constant, or acid dissociation constant, K_a, as defined in equation 2.3 (bracketed values denote their molar concentration). Strong acids have a $K_a > 1$; the stronger the acid, the larger its K_a value and the equilibrium lies further to the right of equation 2.2. The acidity constant varies over many degrees of magnitude and is usually reported in terms of its logarithm pK_a (equation 2.4).

[4]The theory was defined independently by Johannes Brønsted and Martin Lowry in 1923

$$K_a = \frac{[H_3O^+].[A^-]}{[HA]} \tag{2.3}$$

$$pK_a = -log_{10}K_a \tag{2.4}$$

The definition of pK_a mirrors the mathematical relationship between $[H^+]$ and pH (equation 2.1). Indeed, pH can be expressed as a measure of acidity, as shown in equation 2.5. This is called the Henderson-Hasselbach equation and can be used to estimate the pH of a buffer solution.

$$\begin{aligned} pH &= pK_a + log_{10}\frac{[A-]}{[HA]} \\ &= pK_a + log_{10}\frac{[base]}{[acid]} \end{aligned} \tag{2.5}$$

Care must be taken with equation 2.5 for strong acids and alkalis or in dilute solutions (<1mM) because it neglects the dissociation of the acid, the hydrolysis of the base and the ionization of water in a buffer. However, for pK_a values in the range of 5 to 9 the difference between exact calculations, which take these factors into account, and approximate values based on equation 2.5 is less than 5% [68].

2.4.2 Buffers

When a solution consisting of a weak acid and its conjugate base has a pK_a that matches its pH, it is said to have reached its maximum buffering capacity. A buffer is a solution that resists changes in pH, based on the equilibrium set up in equation 2.2. It compensates for any hydrogen ions brought in or removed when small amounts of acid or base are added. It consists of a conjugate acid-base pair so it can donate and accept hydrogen ions to retain the balance required for the target pH. The range over which the buffer is most effective is around $pH = pK_a \pm 1$. Buffer solutions are essential for biochemical processes, both industrially and biologically. Many enzymes, in particular, require precise conditions to work correctly. Their activity may slow or stop and they may even denature if the pH becomes to extreme.

The majority of experiments with ocr were carried out at pH 8 with a buffer solution of Tris-HCl. Tris(hydroxymethyl)aminomethane has a $pK_a = 8.06$ at 25°C and provides effective buffering in the pH range 7.2 to 9.0. It has a molecular weight of

121.1 and a molecular formula of $NH_2C(CH_2OH)_3$ (figure 2.9). The required pH value is obtained by mixing Tris with hydrochloric acid in the proportions shown in table 2.2.

Tris-HCL meets a number of important criteria for buffers [69]:

- Buffer should be highly soluble in water
- Buffer does not absorb in the visible or UV region
- Buffer does not interfere with biological processes
- Buffering capacity should be resilient to changes in salt content
- Buffering capacity should be optimal at the required pH range

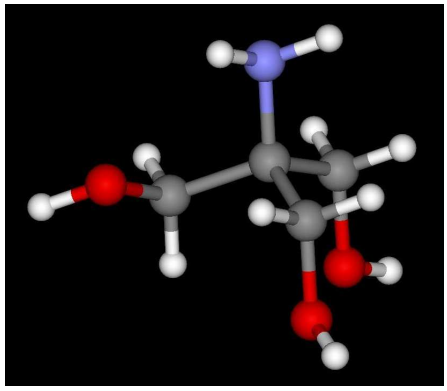


Figure 2.9: Molecular structure of Tris(hydroxymethyl)aminomethane. Data from Chemexper SPRL (*http://www.chemexper.com/*) plotted with Accelrys ViewerLite 5.0 (*http://www.accelrys.com*). Elements are coloured thus: nitrogen - blue, carbon - grey, oxygen - red, hydrogen - white.

Tris-HCl buffer protocol:

ml of HCl	44.2	41.4	38.4	32.5	21.9	12.2	5.0
pH	7.2	7.4	7.6	7.8	8.2	8.6	9.0

Table 2.2: Method to prepare a 25mM Tris-HCl buffer with a specific pH: take 50 ml of 0.1M Tris, mix the appropriate volume of 0.1M hydrochloric acid for the required pH (as stated in the above table) and adjust the final volume to 200 ml with deionized water. Data taken from [69].

2.5 Screening and precipitation

The charge on a protein and the interaction between proteins in a buffer solution is greatly affected by the ionic strength of the solution and the presence of salt ions. Repulsive electrostatic interactions between charged proteins are reduced by even relatively low concentrations of salt ions. This effect is known as screening and is responsible for stabilising (salting-in), and eventually precipitating (salting-out), protein solutions. As the salt concentration is raised, the electrostatic repulsion is increasingly diluted and attractive forces become the dominant interaction. At high enough salt concentration, preferential hydration of salt ions results in water molecules dissociating from the protein, desolvating it. Exposed patches interact, causing the proteins to aggregate, precipitate or crystallise.

2.5.0.1 Screening

When a colloidal particle is dispersed in a polar liquid any ionisable surface groups it possesses dissociate from the particle. It now carries an overall electrostatic charge and is known as a macroion. The oppositely charged counterions from the surface, and any already present in the electrolyte solution, surround the macroion in a diffuse ion cloud known as the electrical double layer (see figure 2.10) [70, 71, 72, 73]. The counterions are held in the vicinity of the macroion by its electric field while diffusing by Brownian motion. Similarly ions with the same polarity as the macroion are repelled from the surface and are known as co-ions. The ion clouds of adjacent macroions overlap when they come too close and the particles are repelled. This helps to avoid aggregation and the suspension is said to be charged stabilised.

If more mobile charges are added, in the form of salt for instance, more counterions gather around each macroion. They form an electrostatic shield that screens the charge on the surface from the surrounding particles. The effective charge in the vicinity of the macroion is reduced, along with the electrostatic repulsion felt by a particle at a distance r. The electrostatic potential for a spherical macroion of radius R and surface potential $\psi_\circ$ decays exponentially as

$$\psi(r) = \psi_\circ \frac{R}{r} e^{-\kappa(r-R)} \tag{2.6}$$

The Debye screening length $1/\kappa$ is a measure of how combinations of ion va-

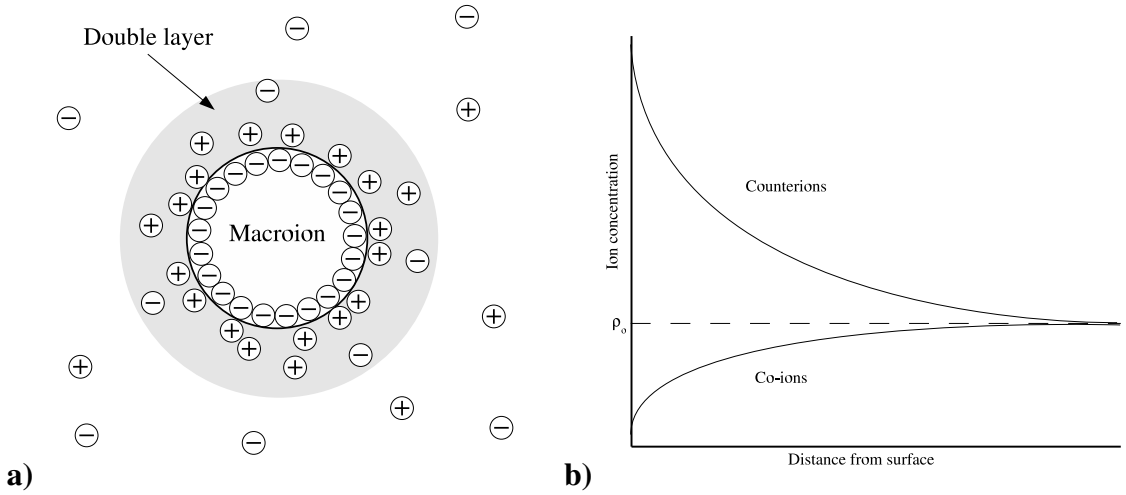


Figure 2.10: **a)** Diagram of a macroion with its diffuse double layer. **b)** Schematic of the charge density around a macroion. There is an accumulation near the charged surface of counterions and a depletion of co-ions. $\rho_\circ$ represents the bulk electrolyte concentration.

lency and concentration contribute to the screening of interactions between charges in solution. It represents the characteristic thickness of the diffuse double layer in figure 2.10 and the distance over which the potential decreases by an exponential factor ($\psi(\frac{1}{\kappa}) = \frac{\psi_\circ}{e}$). The Debye screening length is defined as

$$\frac{1}{\kappa} = \left(\frac{\varepsilon\varepsilon_\circ k_B T}{e^2 \sum c_i z_i^2} \right)^{\frac{1}{2}} \tag{2.7}$$

where $\varepsilon_\circ$ is the permittivity of a vacuum, ε is the dielectric constant of the solvent, k_B is Boltzmann's constant, T is the temperature and e is the electronic charge. The sum is taken over all the counterions present; each ion i has a valency z_i and concentration c_i. For a salt and water solution, where $\varepsilon = 81$, at 25°C the Debye length can be approximated as

$$\frac{1}{\kappa} = \begin{cases} 0.304/\sqrt{[NaCl]} & \text{nm for 1:1 electrolytes (e.g. NaCl)} \\ 0.176/\sqrt{[CaCl_2]} & \text{nm for 2:1 electrolytes (e.g. } CaCl_2\text{)} \\ 0.152/\sqrt{[MgSO_4]} & \text{nm for 2:2 electrolytes (e.g. } MgSO_4\text{)} \end{cases} \tag{2.8}$$

for the molar concentrations of the salts in square brackets. The most important influence on the screening length is the ionic strength, I, of the buffer solution and added

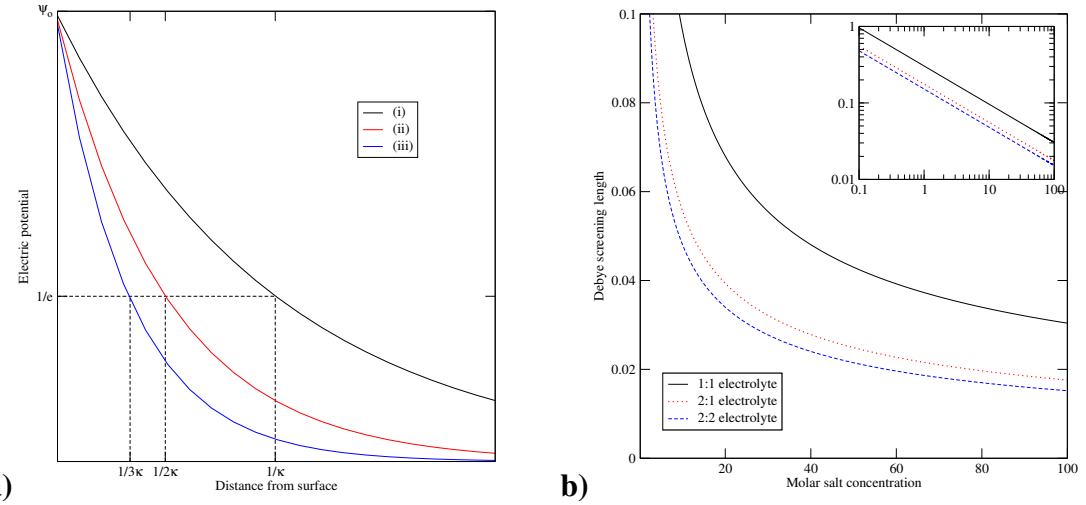


Figure 2.11: Schematics of **a)** variation of electric potential from the surface of a macroion for increasing (i $\rightarrow$ iii) ionic strengths and **b)** variation of Debye screening length with salt concentration for 1:1, 1:2 and 2:2 electrolytes (formula 2.8); inset shows same graph with logarithmic scales.

salt:

$$I = \frac{1}{2}\sum_i c_i z_i^2 \tag{2.9}$$

Figure 2.11 shows the rapid decrease in the range of potential and double layer thickness as ionic strength is increased, either by increasing salt concentration or valency. The electrostatic repulsion between macroions decays in a similar fashion, opening the door to aggregation. The main controlling factors for macroion stability in this situation are the attractive van der Waals potential ($\psi_V(r)$) between two macroions and Coulomb repulsion ($\psi_C(r)$) that arises when their diffuse double layers overlap. This is embodied by the DLVO potential[5], which models the total interaction between pairs of charged colloidal particles as

$$\Psi_{DLVO}(r) = \psi_V(r) + \psi_C(r) + \psi_{HS}(r) \tag{2.10}$$

where $\psi_{HS}(r)$ represents the hard sphere repulsion that prevents the macroions from overlapping. $\psi_{HS}(r) = \infty$ for $r \leq R$, the macroion radius, and $\psi_{HS}(r) = 0$ for $r > R$.

[5]Named after its pioneers: Derjaguin, Landau, Verwey and Overbeek

2.5.0.2 DLVO

Van der Waals forces are short range attractions that arise from transient dipoles generated by the fluctuation of electron clouds. The total interaction potential for identical spherical macroions dispersed in a solvent is found by integrating over all individual pairwise interactions between macroions and is given by equation 2.11 [71].

$$\psi_V(r) = -\frac{A}{12}\left(\frac{4R^2}{r^2-4R^2} + \frac{4R^2}{r^2} + 2ln\left(1-\frac{4R^2}{r^2}\right)\right) \tag{2.11}$$

The Hamaker constant, A, depends on the dielectric properties of the macroions and the intervening solvent. It is typically of the order $10^{-19} - 10^{-20}J$. Equation 2.11 reveals the possibility of irreversible aggregation for particles in close proximity. As $r \rightarrow 2R$ the potential develops a deep minimum at $r = 2R$ (equation 2.12) that can be much greater than the thermal energy k_BT.

$$\lim_{r\rightarrow 2R} \psi_V(r) = -\frac{A}{12}\frac{R}{(r-2R)} \tag{2.12}$$

Charge stabilisation due to the diffuse double layers of macroions provides an opposing force to aggregation. The screened Coulomb repulsion potential from overlapping ion clouds is given by equation 2.13 [71],

$$\psi_C(r) = \psi_\circ \frac{e^{-\kappa(r-2R)}}{(1+\kappa R)^2} \tag{2.13}$$

where $\psi_\circ = \frac{Qe^2}{4\pi\varepsilon_\circ\varepsilon r}$, the potential due to macroions of charge Q. The combination of these repulsive and attractive interactions provides a model potential (equation 2.10) that captures the essential behaviour of a suspension of macroions. The van der Waals potential is largely independent of ionic strength and variations in pH so for a particular solute and solvent it can be considered as fixed. There are a number of situations that may occur, depending on the ionic strength of the electrolyte and surface charge of the macroions. These are summarised below and in figure 2.12.

For highly charged macroions and low ionic strength solvents the large Debye length results in a strong, long-range repulsion and a stable particle distribution. Its peak is known as the energy barrier (see figure 2.12 a). As the ionic strength increases a minimum in the potential develops further out from the energy barrier (figure 2.12 b). The higher the ionic strength the deeper the minimum and a stable equilibrium of macroions can occur under these conditions. The energy barrier decreases for

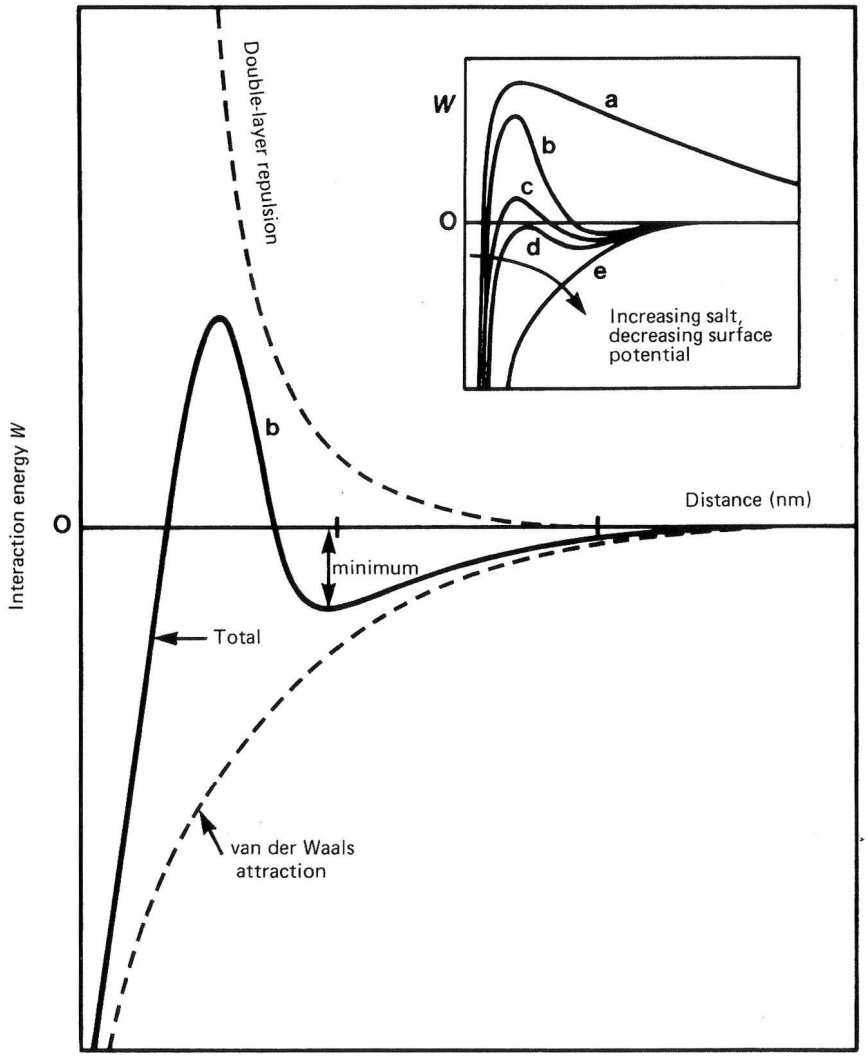


Figure 2.12: Schematic of the DLVO potential as a combination of double layer repulsion and van der Waals attraction for distances from the macroion surface. Scenarios depicted are in order of decreasing Debye length: a) strong repulsion due to highly charged particles and low ionic strength solvents; b) particles exist in stable equilibrium at minimum; c) particles aggregate slowly due to decreasing energy barrier; d) energy barrier falls below axis prompting rapid aggregation; e) charge on particles approaches zero as DLVO potential tends to the pure van der Waals attraction. Figure reproduced from [71].

macroions with a lower charge and a larger electrolyte concentration (see figure 2.12 c). The suspension becomes less stable and the macroions begin to aggregate slowly. Above a particular ionic strength (known as the critical coagulation concentration of the electrolyte) the energy barrier drops below the $\Psi = 0$ axis and the macroions rapidly aggregate (see figure 2.12 d). Any resistance to aggregation is lost as the macroion charge approaches zero. The DLVO potential approaches the van der Waals potential and the attraction is felt at all separations.

This interplay between attractive and repulsive potentials is the essence of the DLVO theory. In particular, it shows that by increasing the salt concentration, even highly charged particles can aggregate. The effective attraction introduced by the salt and its associated degree of screening is controlled by the ionic strength of the solution and the Debye screening length, κ^{-1}.

2.5.1 Precipitation and the Hofmeister series

The solubility of a protein in a salt solution (and therefore crystallisation) is affected by the addition of salts and, furthermore, depends on the type of salt present. Franz Hofmeister experimentally observed the precipitation of proteins in solutions of different salts. He discovered that certain salts have a consistently greater effect, regardless of the protein used [74, 75, 76]. The Hofmeister series is an empirically ordered list of salt anions and cations in terms of their protein precipitating efficiency. There is no quantitative theory of protein solubility with salt type so there are some circumstances where the order is quoted slightly differently but predominately they are displayed as in table 2.3. The anion typically has the greatest influence, while the effects of cations are usually considered independent and additive.

At the interface between proteins and solvent, interactions like electrostatics and solvent surface tension significantly affect protein solubility. Salts also interact preferentially with water molecules and can affect their structure. In a low ionic strength solution the screened proteins have a lower electrostatic free energy and a greater solubility. In this way low concentrations of salts, regardless of their nature, tend to stabilise the protein in solution. As the salt concentration increases the effect it has depends on its type and where it lies in the Hofmeister series.

The salts near the start of the series decrease the solubility of proteins and they precipitate (“salting-out”). Physically, the salts perturb the structure of water and increase

Anions:	SO_4^{2-}	$\simeq$	F^-	>	HPO_4^{2-}	>	CH_3COO^-	>	Cl^-	>	Br^-
	(salting-out)										*(salting-in)*
Cations:	NH_4^+	>	K^+	>	Na^+	>	Li^+	>	Mg^{2+}	>	Ca^{2+}

Table 2.3: A common depiction of the Hofmeister series [74]. The salt ions that precipitate proteins with the greatest efficiency lie further to the left in the series. The series is not based on a fundamental law but on empirical results that suggest it holds true for the majority of proteins.

its surface tension. Salts at the other end of the series have the opposite affect. The structure of water is reinforced, the surface tension is not affected and the solubility of proteins is increased, stabilising them in solution ("salting-in"). The change between salting-in and salting-out, while gradual, is commonly accepted to occur at Na^+ and Cl^-; high concentrations of NaCl have only a small affect on protein solubility. This is demonstrated in figure 2.13, which shows the solubility of hemoglobin in various electrolytes. Electrostatics is thought to be the most significant factor in describing the behaviour of salt and protein interactions. Other factors must still be considered though, including ion size and the interaction between water molecules and polarizable ions. Salts also interact directly with proteins, potentially binding to residues or denaturing the protein. For instance, although some divalent cation salts like $CaCl_2$ and $MgCl_2$ raise the surface tension of water, they do not reduce protein solubility because they bind to the proteins. Bostroöm *et al.* have also suggested that the more polarizable an ion, the stronger its attraction to the protein surface and the more effective it is at screening [76].

The Hofmeister series has been encountered in numerous other phenomena. For example, it also reflects the increased ability of some salts to aid in the stability of protein conformations and protein crystallisation. Ammonium sulphate is particularly popular, not just for its prominent position in the series but because it is highly soluble, cheap and typically has no permanent effect on protein conformation.

2.6 Protein charge and titration curves

The charge on a protein is highly sensitive to its local environment but can be maintained at a constant value by fixing the pH of the buffer solution. It is therefore impor-

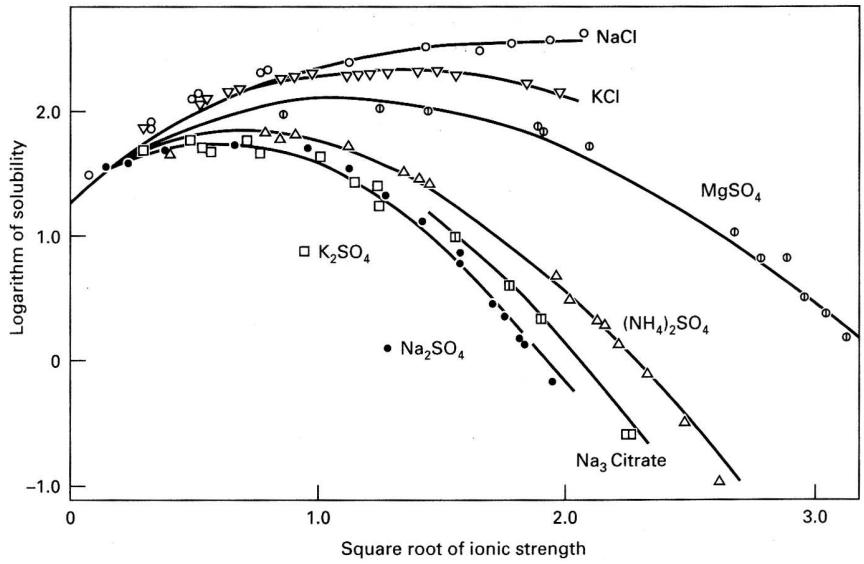


Figure 2.13: Diagram showing the solubility of hemoglobin in a range of electrolyte types and concentrations. Solubility is measured in grams per kilograms H_2O. Reproduced from [74].

tant to know how the charge varies with pH. This can be determined experimentally, by measuring the hydrogen-ion activity [77], or theoretically from the intrinsic properties of the protein's constituent amino acids. At a given pH the ionisation state of each titratable group in a protein depends on their individual pK_a values. Four out of the twenty amino acids have pK_a values that mean they are fully charged at neutral pH. The pK_a's of the highly acidic residues aspartic and glutamic acid are much lower than 7, so both should contribute -1 to the charge of a protein. Conversely, the pK_a's of the highly basic residues lysine and arginine are much greater than 7, so they should contribute +1 to the protein charge. However, although the pK_a values for individual amino acids are known (see [78] and table 2.4) combining them to determine the overall pK_a of the protein is non-trivial.

Experimental titration curves for a native protein often deviate significantly from those based on the sum of the intrinsic properties of its individual groups. This simple theoretical prediction is only valid for a denatured protein with no interactions between groups, and does not take into account the ionic strength of the electrolyte. The most substantial difference is attributed to electrostatic interactions between titratable groups, which alters each one's protonation equilibrium (equation 2.2) [79]. This

```
KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS
TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS
DGNGMNAWVA WRNRCKGTDV QAWIRGCRL
```

Figure 2.14: 129 amino acid sequence of hen egg-white lysozyme [80, 81]. The amino acid that each letter represents can be found in table 2.1.

is especially noticeable when the ionic strength is increased, and the difference between experimental and theoretical curves is reduced (an effect discussed below). The three dimensional structure of a protein may also have a significant effect. Titratable groups may be buried in a hydrophobic environment, which restricts access for protonation or deprotonation events and reduces those groups' effect on the overall pK_a compared to exposed residues. Alternatively, there may be hydrogen bonds between surface groups, which can perturb their pK_a values. More advanced models attempt to take into account protein structure and the fact that discrete charges are spread non-uniformly over the molecule. They commonly describe each titratable group by its intrinsic value when isolated, plus a perturbation that represents all the effects (electrostatic, hydrophobic etc) of containing the group within a native protein molecule:

$$pK_a = pK_{a,intrinsic} + \Delta pK_a \tag{2.14}$$

However, the magnitude of these effects is hard to predict and, as such, hard to incorporate. Thus, while theoretical treatments to determine a protein's pK_a are possible they should be compared, where available, to experimental results.

2.6.1 Lysozyme

A useful illustration of the variation in titration curves can be seen with the protein lysozyme. Lysozyme is commonly used as a generic protein and is well characterized. The primary structure of lysozyme is shown in figure 2.14. It is a globular protein with a molecular mass of 14313 Daltons and an average radius of 1.6nm. At pH 8 it has a charge of around +8e and it has a pI of 10.7. Its native structure is shown in figure 2.15.

Kuehner et al. measured the titration curve of lysozyme in a potassium chloride solution of varying concentration (from 0.1M - 2M) [77]. They also compared the 0.2M curve to both experimental and theoretical curves determined by other authors.

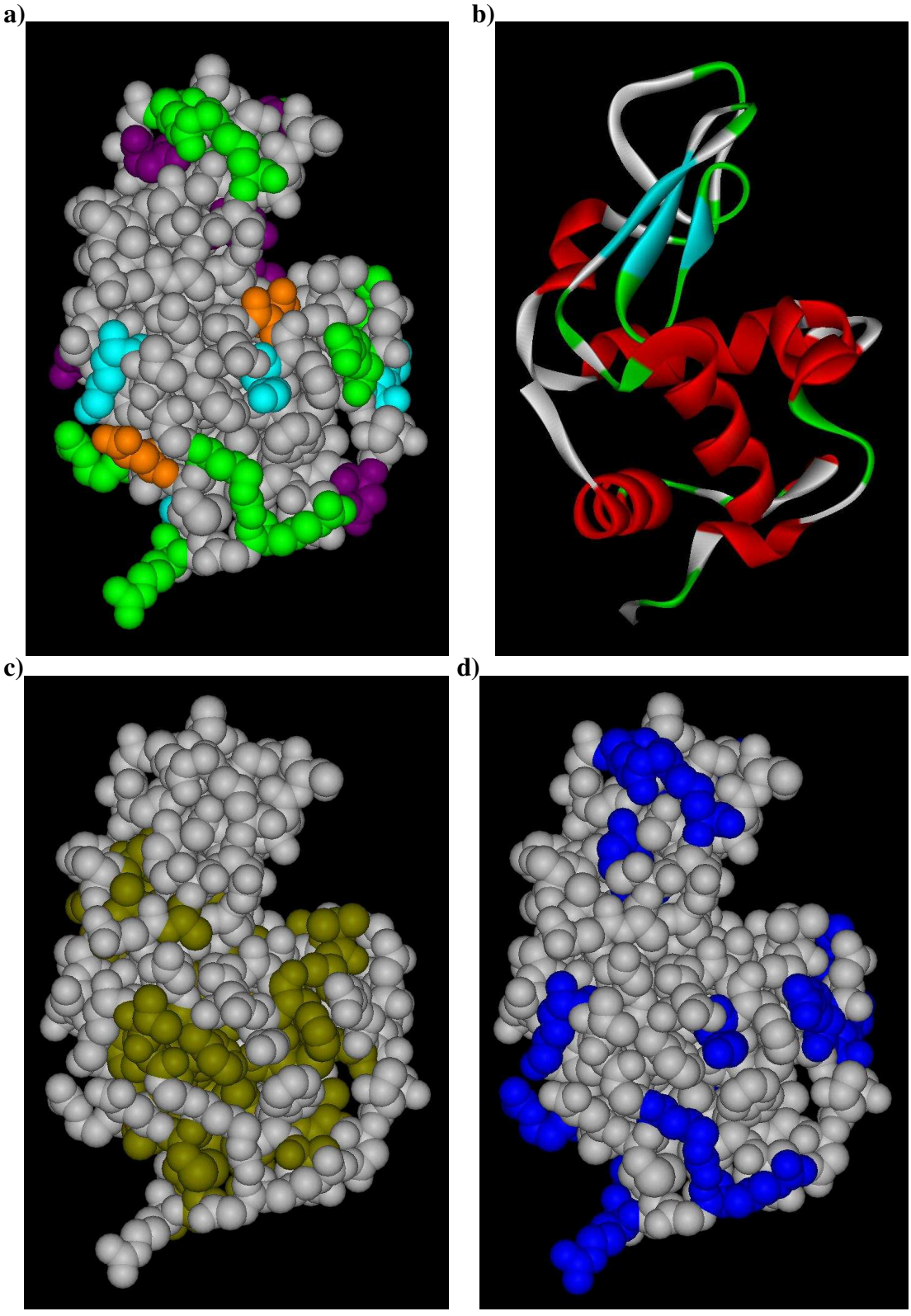


Figure 2.15: The native structure of lysozyme. **a)** Charged amino acids: purple = aspartic acid (-), orange = glutamic acid (-), blue = lysine (+) and green = arginine (+). **b)** Secondary structure elements: red = α-helices, blue = β-sheets. **c)** Hydrophobic groups **d)** Basic groups. As the ionic strength of the solution increases the charge is screened out and the hydrophobic sticky patches are revealed, prompting protein aggregation. Molecules plotted with Accelrys ViewerLite 5.0 (http://www.accelrys.com)

.

Their results can be seen in figure 2.16. The theoretical curve (based on equation 2.14) differs from the measured results below pH 4 and above pH 11 by up to +1 charge unit but in between agrees very well. Salt has the most effect below pH 4.3 where the titration curves are clearly separated and indicate an increasing net charge on the protein with increasing ionic strength. Above pH 4.3 the curves converge to within 0.5 charge units. The 2M curve differs slightly, showing the lowest charge in the pH range 5.2 to 9.1. However, some aggregates of protein were observed in the 2M solution, which may have affected its behaviour during the titration.

2.6.2 Ocr

A number of protein and DNA theoretical tools can be found on the Expasy website (*http://ca.expasy.org/tools*). An on-line program for calculating the theoretical titration curve for any protein is provided by the IUT de Provence, site d'Arles (*http://www.iut-arles.up.univ-mrs.fr/w3bb/d_abim/compo-p.html*). This uses the protein's primary sequence and the pK_a values for the amino acids shown in table 2.4 to calculate the charge at various pH values. The titration curve predicted is shown by the black line in figure 2.17a). The ocr dimer has a pI of 3.67 and a charge of $-56e$ at pH 8. No experimental titration curve for ocr is available but titration curves based on an advanced theoretical model that takes into account the ionic strength were kindly provided by Professor Bo Jönsson of Lund University (figure 2.17). The protein is represented by a linear chain of monomer titrating sites that can only interact through a screened Coulomb potential [82, 83]. A Monte Carlo simulation is run using a grand canonical ensemble. A random monomer is chosen and its charge altered, resulting in a change in free energy, ΔE, that stems from the usual acid-base reaction of an isolated monomer and the screened Coulomb interactions between monomers, ΔE_C:

$$\Delta E = \Delta E_C \pm k_B T \ln 10 (pH - pK_{a,0}) \tag{2.15}$$

where the plus and minus sign represent protonation and deprotonation events respectively. The $pK_{a,0}$ is the intrinsic dissociation constant of the monomer, which will be lower than the apparent pK_a due to electrostatic interactions between charged monomers reducing the dissociation of the protons. The difference between the apparent and intrinsic pK_a is related to the excess free energy, E, by

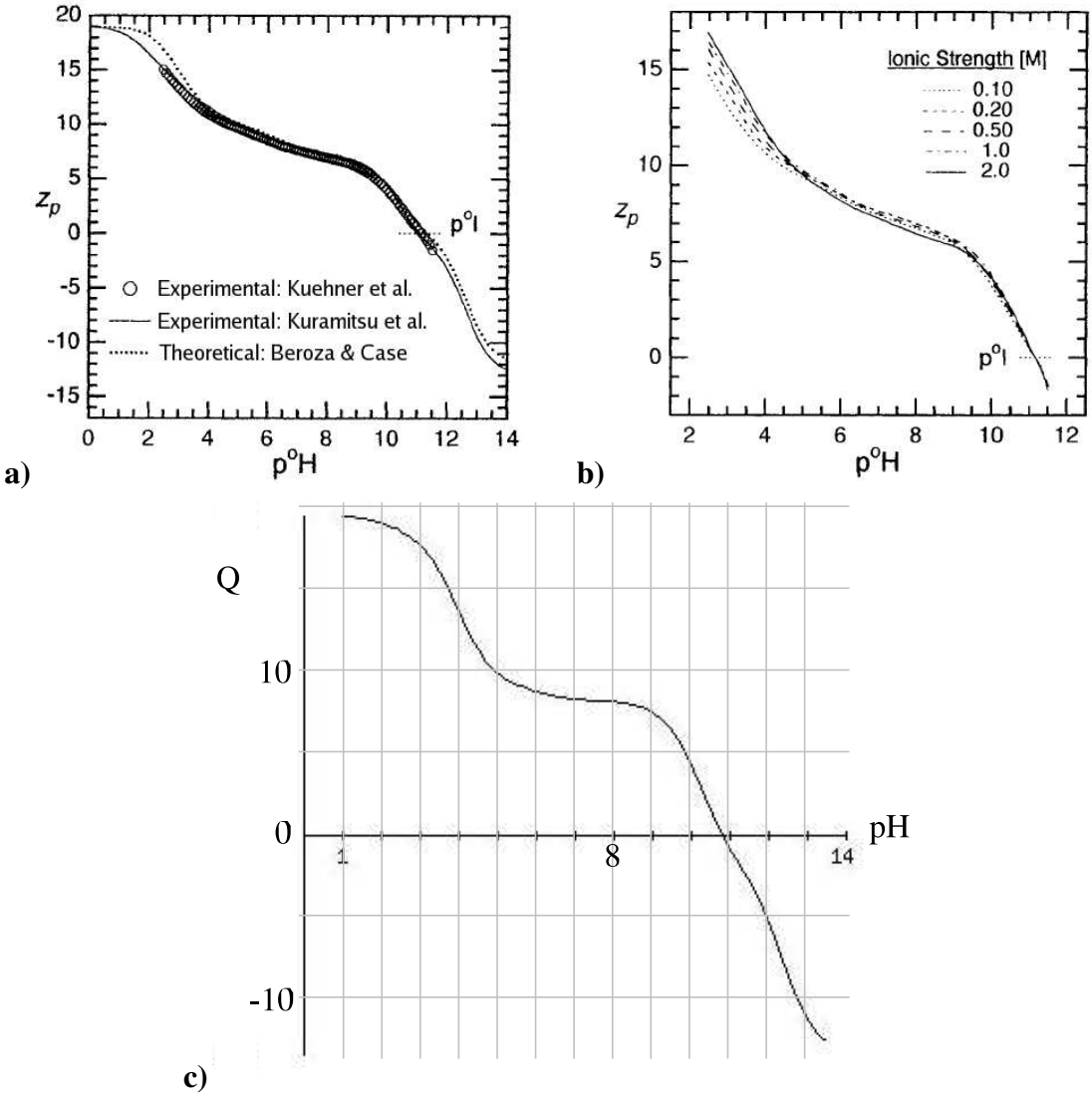


Figure 2.16: **a)** Experimental and theoretical titration curves for the protein lysozyme in 0.2M potassium chloride. **b)** Experimental titration curves for lysozyme in varying concentrations of potassium chloride at 25°C. Diagrams reproduced from [77]. **c)** Theoretical titration curved based on primary structure of lysozyme (calculated by on-line program *http://www.iut-arles.up.univ-mrs.fr/w3bb/d_abim/compo-p.html*).

$$\Delta pK_a = pK_a - pK_{a,0} = \frac{1}{k_B T \ln 10} \frac{\partial E}{\partial \alpha} \tag{2.16}$$

Ocr is predominantly negative due to the high number of aspartic and glutamic acid residues, most of which are spread over the surface of the protein. They have low pK_a values and deprotonate, becoming negatively charged, as the pH increases. However, each feels an electrostatic repulsion from its similarly charged neighbours. This encourages protonation of the aspartic and glutamic acid groups, reducing the net negative charge on the protein. Such an effect is observed in the simulations of Bo Jönsson (figure 2.17a) where the charge of ocr above its pI is less negative than the one predicted using only its primary structure. The effect of increasing ionic strength is to screen the charged residues from each other and consequently reduces the protonation induced by inter-residue repulsion. The perturbations in pK_a relax and, as observed in figure 2.17b), the charge on ocr becomes more negative. The titration curve taking into account inter-residue interactions approaches that of the titration curve based on primary structure with increasing ionic strength. While electrostatics is the most dominant effect, the curves do not meet due to other non-electrostatic interactions and structural properties, like buried residues. Ocr has various titratable groups that are partly buried and at least four (two tyrosine and two arginine) that are fully buried. The acidic, basic, hydrophobic and hydrophilic groups of ocr are shown in figure 2.18.

Ocr not only has a high charge, it has a high surface charge density. An ellipsoid with radii a and c has a surface area given by $A = 2\pi(a^2 + c^2 \frac{\alpha}{tan\alpha})$, where $\alpha = cos^{-1}(\frac{a}{c})$. This gives 68nm^2 for ocr, which therefore has a surface charge density of -0.82e/nm^2 (for Q=-56 at pH 8). In comparison, lysozyme can be modelled as a sphere with a surface area of 32nm^2 and a surface charge density of 0.25e/nm^2 (for Q=8 at pH 8), over three times lower than that of ocr. To put the charge of ocr and lysozyme in context we look in the next section at the entire set of proteins expressed by *E.coli*, known as its proteome. Ocr is shown to have extreme properties shared with only a handful of other proteins.

2.6.3 E.coli proteins

Most bacteria have a few thousand proteins, a large proportion of which are neutral. The rest typically form a Gaussian distribution centred near zero. This is shown in figure 2.19 for the *E.coli* K-12 organism studied by Sear [84]. This strain has 4358

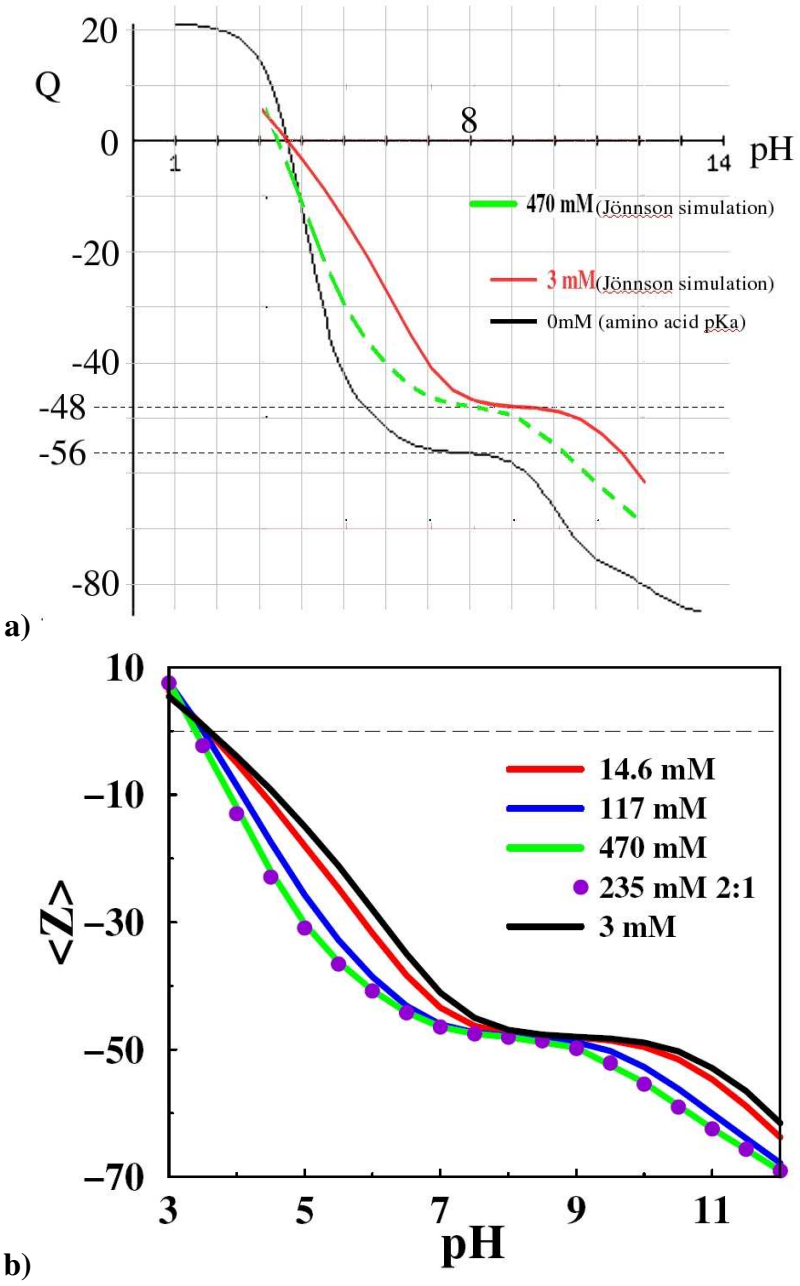


Figure 2.17: Titration curves for ocr. **a)** The black line (0mM) represents the theoretical titration curve of dimeric ocr based on the pK_a of each titratable group in its primary sequence. It predicts that ocr has a pI of 3.67 and a charge of -56 at pH 8.It was calculated with the on-line program at *http://www.iut-arles.up.univ-mrs.fr/w3bb/d_abim/compo-p.html*. The red (3mM) and green (470mM) lines are taken from graph **b)** for comparison. **b)** Monte carlo simulation titration curve of ocr in varying molar concentrations of salt, including a 2:1 electrolyte representing $(NH_4)_2SO_4$, courtesy of Bo Jönsson. It predicts that ocr has a pI of 3.6 at 3mM salt and a charge of -48 at pH 8. Increasing the ionic strength of the solution increases the negative charge of ocr for pH values above its pI.

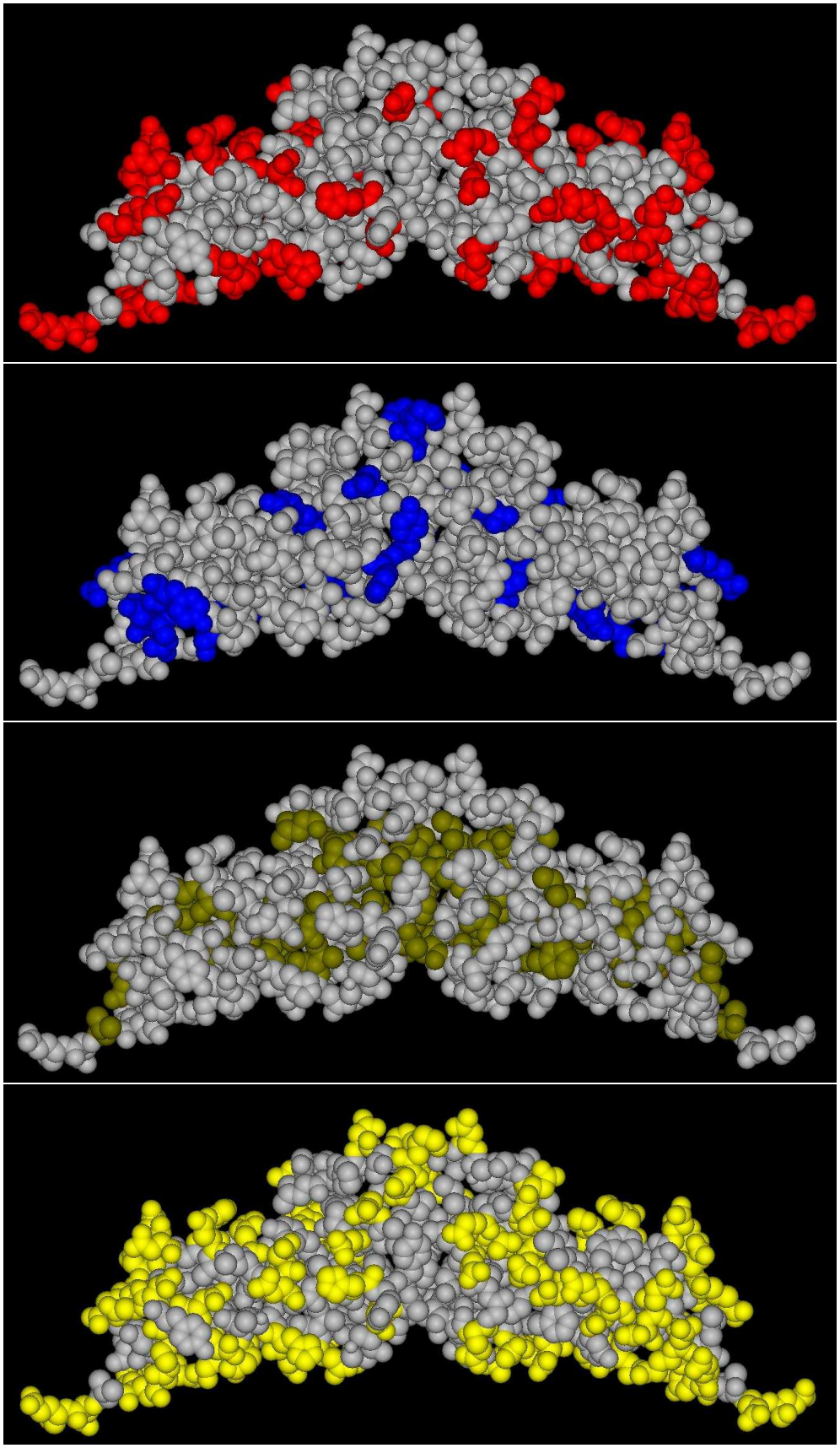


Figure 2.18: From top to bottom: acidic (red), basic (blue), hydrophobic (brown) and hydrophilic (yellow) groups on ocr.

Residue	pK_a		Residue	pK_a
α-NH_3	9.69		α-COOH	2.34
Arginine	12.40		Aspartic acid	3.86
Lysine	10.50		Glutamic acid	4.25
Histidine	6.00		Tyrosine	10.00

Table 2.4: The pK_a values used by the IUT de Provence d'Arles on-line program to calculate the theoretical titration curve for a protein from its primary sequence. *[http://www.iut-arles.up.univ-mrs.fr/w3bb/d_abim/compo-p.html]*

proteins. The charge on each protein at pH 7 was estimated simply by subtracting the total number of its constituent aspartic and glutamic acids from the total number of its constituent lysine and arginine amino acids. This assumes only those four amino acids have pK_a values significantly greater or less than 7 (as described above). The Gaussian fitted by Sear is roughly symmetric, centred at the mean charge of -2.16 and has a width of the order 10.

The charge on a protein naturally varies with its size. A scatter plot of the charge Q and size, or number of constituent amino acids, M of each protein in *E.coli* is given in figure 2.20a). The data was kindly provided by R.Sear from his work with *E.coli* K-12. A wide range of charges are possible for any particular protein size but the larger proteins are on average more highly charged. Data points for ocr and lysozyme have been added to the plot for comparison. There are a few *E.coli* proteins with charges of -50 or below but they are all significantly larger than ocr (from two to seven times the number of residues). Six of them are also hypothetical proteins. A hypothetical protein is one whose existence has been predicted during genome sequence analysis but lacks experimental evidence. Lysozyme, on the other hand, has a charge of 8 and is firmly located alongside the main bulk of *E.coli* proteins.

Assuming the volume and mass of a protein scales with M, estimates can be made for charge densities on proteins. Ocr has a charge to size ratio of $Q/M = -0.239$ e/residue, over ten times the average ratio of -0.01 e/residue. There are only five proteins with comparable ratios, which range from -0.195 to -0.297 e/residue. Lysozyme has a ratio of 0.062 e/residue, once again agreeing with the average protein properties of *E.coli*. If instead we look at $\frac{Q}{M^{2/3}}$ we gain a measure of the surface charge density (assuming all charges lie on the surface). Figure 2.20b) shows a graph of Q against

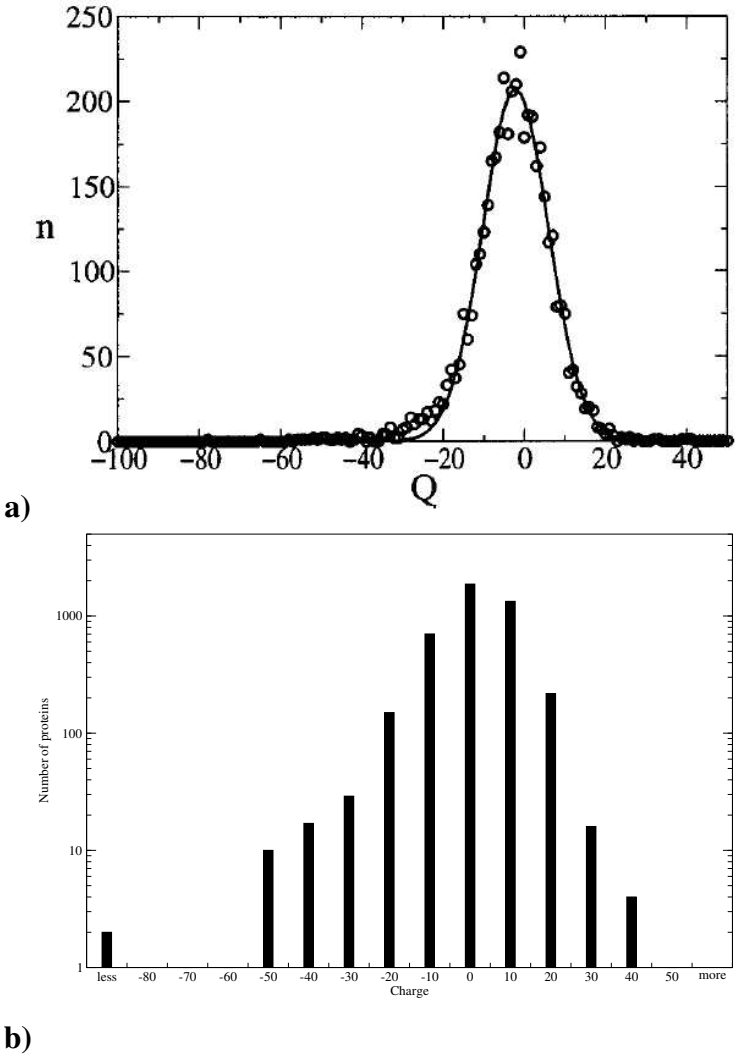


Figure 2.19: The number of *E.coli* K-12 proteins with a net charge Q. Two proteins with charges less than -100 are not shown. The data is fitted to a Gaussian curve in a) and plotted logarithmically as a histogram in b). [Figure a) reproduced from [84], figure b) produced from data kindly provided by R.Sear.]

$\frac{1}{M^{2/3}}$ for ocr, lysozyme and all the proteins in *E.coli*. Ocr has a surface charge density of -1.47, while lysozyme has the unremarkable 0.31. Only one protein has a surface charge density larger than ocr and it is a hypothetical protein with -1.53. The three most comparable proteins have surface charge densities of -1.33, -1.01 and -1.00. The first is the acidic protein msyB but the other two are hypothetical proteins. These results highlight ocr's position as not only a highly charged protein but one with a very large surface charge density, lying well outside the average properties for the majority of proteins.

2.6.4 Ocr precipitation

Proteins exist in cells at salt concentrations of the order 0.15-0.3M and are very soluble under these physiological conditions. While the conditions for protein precipitation depend on the identity of both the salt and protein, as discussed above, it typically requires salt concentrations of less than 2M ammonium sulphate. Ocr, however, does not follow this trend. It remains in solution for all salts studied up to saturation, except ammonium sulphate. For ammonium sulphate it remains stable to 3.2M, where it forms a transparent gel (figure 5.3). Understanding this strange behaviour of ocr is the main subject of this thesis. The transition is rapid and is discussed in section 5.1. The crucial relationship between charge and pH, described above, has important consequences for ocr precipitation. It is investigated in section 5.4 and discussed in section 7.3.

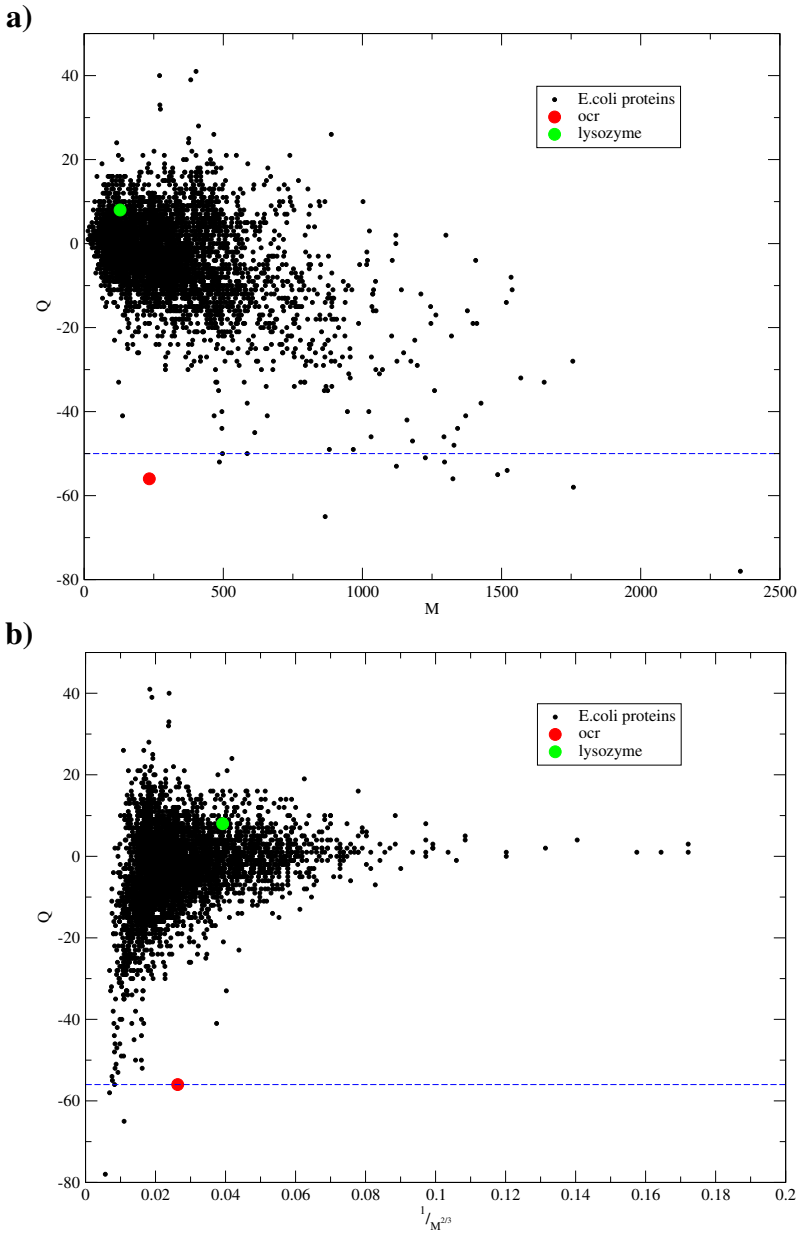


Figure 2.20: Scatter plots showing the charge of ocr, lysozyme and each protein in *E.coli* K-12 against a) the number of amino acids, M and b) $\frac{1}{M^{2/3}}$, giving a measure of surface charge density. Ocr has properties far outside those of the average protein. [Data kindly provided by R.Sear.]

Chapter 3

Expression of ocr

3.1 Protein production

Proteins are synthesized in a cell by a process called expression. Expression follows the central dogma of molecular biology, where DNA (deoxyribonucleic acid) is transcribed into RNA (ribonucleic acid), which in turn is translated into proteins. DNA is composed from the nucleotides adenine (A), guanine (G), cytosine (C) and thymine (T), which form two complementary strands twisted into a double-helix conformation. Each nucleotide in one strand has a complementary partner in the other, to which it hydrogen bonds: adenine with thymine and guanine with cytosine. RNA is also composed of nucleotides but forms a single strand in which thymine is replaced by a fifth nucleotide, uracil (U). A section of DNA that encodes (contains information) for RNA or proteins is called a gene.

During prokaryotic expression a gene's DNA sequence is converted, via RNA , into the polypeptide sequence of amino acids that make up the protein's primary structure. This natural process has two main steps: the transcription of DNA and the translation of mRNA [85]. Before a gene's coded information can be used it is transferred, or transcribed, into mRNA, a single stranded copy of the DNA sequence called messenger RNA. Messenger RNA (mRNA) acts as a template for the synthesis of proteins, specifying the required sequence of amino acids. The nucleotides on mRNA form a code where each three letter sequence, called a codon (eg. CAG), refers to a certain amino acid. The mRNA carries this information to one of the cell's ribosomes, which translates the instructions and builds the protein chains. The ribosome matches each

codon sequence to the sequence of a tRNA molecule, which carries the required amino acid. As the amino acids are being added to the growing polypeptide chain it spontaneously begins curling into its secondary structure elements. These, in turn, fold around each other to form the tertiary structure that gives the protein its characteristic, functional shape known as its native conformation.

Natural protein expression can be utilised to produce specific proteins. A preliminary step called transformation is introduced where foreign strands of DNA are inserted into bacteria to create a strain suitable for producing the required protein. A circular piece of DNA called a plasmid is used. It contains the gene that holds the information needed to produce the protein and a gene for antibiotic resistance. One gene encodes the ocr protein, the other causes an *E. coli* cell to become resistant to the antibiotic carbenicillin, which is usually lethal to bacteria. By growing cells in the presence of the antibiotic we can apply a selection pressure to encourage the cells best suited to ocr production to grow. Any surviving cells have been transformed and contain the gene required for expressing the protein. Once transformation is complete the plasmid DNA is transcribed and translated into ocr proteins by the cells as if it were a native gene.

Methods for protein expression and the purification steps needed to remove unwanted by-products [86] are described below. They are based on those employed by Studier, Walkinshaw, Atanasiu, Sturrock *et al.* [61, 65, 64, 55, 66, 62, 60, 87] and have been refined through practice. Two essential techniques for analysing the protein purification process are gel electrophoresis and near UV absorbance spectroscopy.

3.1.1 Gel electrophoresis

Gel electrophoresis is a technique for separating proteins according to their electrophoretic mobility, or movement under electric field, through a polymer network. Electrophoretic mobility depends chiefly on protein molecular weight, 3-dimensional shape and charge. A macromolecule of charge q and characteristic size l undergoing electrophoresis experiences two opposing forces due to an electric field E, $F = qE$, and a viscous drag $F_d = c\eta lv$, where v is the velocity of the macromolecule, η the viscosity of the medium and c is a constant of proportionality [88, 89]. Since the charge depends on the size of the macromolecule both forces scale in an similar manner with size. When these forces are balanced a terminal velocity is reached that depends predominantly on the

particle's charge: $v = qE/c\eta l$. If a porous gel is used in place of a free solution, the macromolecules must filter through a network of randomly shaped pores and are additionally separated on the basis of 3-dimensional size. SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) is a refinement whereby separation can be achieved solely on the basis of molecular weight [86, 90].

SDS is an ionic surfactant that denatures the protein by breaking the non-covalent interactions between its secondary and tertiary structures. Proteins are then separated during electrophoresis by the number of amino acids in their polypeptide chain (ie. molecular weight), rather than physical shape. Furthermore, one SDS anion binds to the peptide chain for every two amino acid residues (~1.4g SDS/1.0g protein), conveying an overall negative charge to the protein proportional to the number of amino acids. Since the resulting denatured proteins all have a similar charge to mass ratio, electrophoretic separation is determined by molecular weight alone.

The gel used to separate the proteins is a crosslinked polymer, usually polyacrylamide. Choosing the porosity, or mesh size, of the gel depends on the weight of the protein being investigated. It is controlled by the concentrations of polymer and cross-linker. 15% polyacrylamide gels were used for experiments involving ocr. A full description of the method for making gels, including protocols for the required buffers, is given in appendix A. The procedure is based on the Laemmli method of one-dimensional, denaturing, discontinuous gel electrophoresis [91]. A gel is made from acrylamide cross-linking initiated by APS[1] and TEMED[2] (a catalyst that increases the rate of polymerisation). The "discontinuous" gel refers to a small layer of highly porous "stacking" gel, in which the proteins are initially loaded, above the main "separating" gel. At the start of electrophoresis there is an initial ion gradient that concentrates all the proteins into a small area. The large pores in the stacking gel ensure no separation occurs before the proteins enter the main gel, where they are separated by the sieve-like nature of the cross-linked polyacrylamide.

A number of different samples can be tested at once. Each is put into its own well in the stacking gel and will travel down the gel in parallel, individual columns. The proteins in a sample are separated vertically in the column, each appearing as a small horizontal band once stained. Proteins travel a distance that is approximately inversely proportional to the logarithm of their size. Thus bands at the same height in

[1] ammonium persulphate
[2] tetramethylethylenediamine

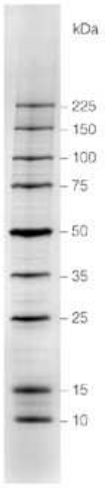


Figure 3.1: A broad range protein molecular weight marker from Promega. It consists of nine clearly identifiable bands for molecular weights varying between 10kDa to 225kDa. Image reproduced from Promega instruction leaflet.

different columns correspond to proteins of the same size, or indeed the same protein. A molecular weight marker and a control sample of pure ocr were also run on the gel to provide a reference for protein mobility. A marker contains different proteins with a range of known weights. They appear on an electrophoresis gel as a ladder that can be used to approximate the size of unknown proteins. An example of a broad range protein molecular weight marker from Promega is shown in figure 3.1. It has nine evenly spaced bands for protein sizes ranging from 10kDa to 225kDa. When analyzing gels care should be taken with proteins of similar weight as they will co-migrate. For example the bands for lysozyme (14400 Da) and ocr (13678 Da) will appear close together on the gel.

Electrophoresis was performed with a Mighty Small II (Hoefer Scientific instruments) vertical slab gel kit at 25-30mA per gel until the smallest proteins have reached the bottom of the gel (~1 hour). The proteins in the gel are made visible by staining them with coomassie blue dye. Gels are placed in a solution of dye, called “stain”, and gently shaken for around 2 hours. To remove excess dye that has not attached to proteins, the gels were then placed in a destain solution of methanol and acetic acid and gently shaken until clear protein bands can be seen (~3-10 hours). Electrophoresis gels can be seen in figures 3.3 and 3.5.

3.1.2 Near UV absorbance spectroscopy

Protein concentration in a solution can be calculated by measuring the amount of UV radiation it absorbs. At near UV wavelengths (250nm-300nm) protein absorption is predominantly due to tryptophan and tyrosine amino acids. Both absorb strongly in this range and derive their UV spectral characteristics from their aromatic sidechains. The combined effect of these residues means that most proteins exhibit a broad absorbance spectrum with a peak maximum at around 280nm [53].

When electromagnetic radiation is incident on a molecule it may induce electronic transitions where the valence electrons absorb the energy of the photons and are excited from their ground state ($E_\circ$) to a higher energy level (E_n). A photon of wavelength λ is absorbed by a molecule if its energy ($\frac{hc}{\lambda}$) matches the difference between these energy levels ΔE:

$$E_n - E_\circ = \Delta E = \frac{hc}{\lambda}$$

where h is Planck's constant and c is the speed of light [92, 67]. Absorbance spectra of molecules in solution are much broader than the fine structure observed in gas phase absorbance spectra. One of the main reasons for this is the fluctuating local environments experienced by molecules, which results in slightly different energy level spacings. Collisions with neighbouring molecules also lead to broader bands as they reduce the lifetime of the excited state, which affects its energy due to the Heisenberg uncertainty principle $\delta E \delta t \geq \frac{h}{2\pi}$ [93].

An example of a UV spectrum measured with a Hitachi U-2001 spectrophotometer over the range 220-340nm for a 5.4mg/ml solution of ocr is shown in figure 3.6. The absorption due to the buffer (20mM Tris-HCl, pH 8) is accounted for by subtracting the spectrum of a pure buffer solution, called a *baseline*. The protein concentration can be calculated from the absorbance spectrum by using the Beer-Lambert law (equation 3.1):

$$A_\lambda = \varepsilon_\lambda c l \tag{3.1}$$

$$A_\lambda = log_{10}(\frac{I_\circ}{I_t}) \tag{3.2}$$

The absorbance of a sample at a wavelength λ is defined in equation 3.2 where

$I_\circ$ is the intensity of incident light and I_t the intensity of transmitted light. The Beer-Lambert law states that the absorbance is proportional to the molar concentration of protein (or other absorbing molecule) c, the path length or thickness of sample l (usually 1cm) and the molar extinction coefficient (also known as molar absorption coefficient). This latter quantity is a measure of how strongly the molecule interacts with radiation of wavelength λ. It is measured in units of $M^{-1}cm^{-1}$, which can be rewritten in terms of cm^2/mol. So the molar extinction coefficient can be thought of as a molar cross-sectional area for absorption where molecules with larger ε can block (*i.e.* absorb) the incident radiation more effectively.

The molar extinction coefficient is a constant for a particular protein at a specific wavelength in a specific buffer. In proteins different amino acids have different absorbances at 280nm. Tryptophan (W) absorbs the most, closely followed by tyrosine (Y). This is mainly due to the presence of aromatic rings but there is also a small contribution from cysteine (C) pairs due to the formation of disulfide bridges. Phenylalanine, although it contains an aromatic side chain, only absorbs at lower wavelengths (240-265nm) and is an order of magnitude weaker. Furthermore, amino acid residues exposed on the surface will contribute differently to those buried inside the protein structure. The theoretical extinction coefficient for a denatured protein (when all residues are exposed) can be calculated from the number of tryptophan, tyrosine and cysteine amino acids present [94]:

$$\varepsilon_{280} = 5690n_W + 1280n_Y + 120n_C$$

The ocr dimer has 2 tryptophan, 16 tyrosine and no cysteine amino acids, giving a theoretical molar extinction coefficient at $\lambda = 280$nm of 31860 $M^{-1}cm^{-1}$. The molar extinction coefficient of native ocr was calculated by measuring its absorbance and comparing it to the absorbance of denatured ocr at the same concentration, giving a value of 31095 $M^{-1}cm^{-1}$ [87].

The absorbance peak for DNA is centred around 260nm and when DNA is present alongside proteins their spectra overlap. An example of this is shown in figure 3.6. If the ratio of absorbance at 260nm:280nm is above ~ 1.6 then the level of DNA present is acceptably low. If it is less than this, further purification is required to remove the contaminating DNA but the protein concentration (in mg/ml) can still be estimated very roughly using the Warburg-Christian equation [95]:

$$c = 1.76A_{280} - 0.55A_{260} \qquad \text{mg/ml of protein} \tag{3.3}$$

3.2 Culture growth

To produce wild-type ocr protein the *E.coli* strain BL21(DE3)pLysS was transformed with plasmid pAR2993. BL21(DE3) host cells are highly efficient, all-purpose strains that are widely used for gene expression. BL21(DE3) hosts contain T7 RNA polymerase, which naturally carries out small amounts of transcription in uninduced cells [96]. This behaviour is suppressed by T7 lysozyme (which is expressed by the pLysS plasmid) as it is a natural inhibitor of T7 RNA polymerase. So hosts carrying pLysS plasmids provide a tighter control over expression, where pLysS is maintained by chloramphenicol. The natural transcription is greatly enhanced (or induced) when IPTG is added.

The procedure for growing cells and expressing proteins is described below and should be carried out under sterile conditions using materials and media sterilised in an autoclave. The original protocol in the Dryden group was adequate for molecular biologists, requiring only micrograms for their experiments. However, large quantities of ocr were required for the repeated measurements within this thesis. Particularly for rheology, which needs milligrams of ocr. I made a number of improvements to the protocol that increased production from 1-10mg ocr per prep, to over 100mg ocr per prep. The improved procedure also reduces the time needed to perform the protocol by two days. The main changes I introduced were the use of Overnight Express for culture growth and trichloroacetic acid during the purification stage. The original protocol and the improvements made are described in the following sections.

3.2.1 Transformation

Competent cells (100μl) were thawed on ice for 10 minutes (competent means that the cell envelope is permeable to DNA). 1μl of (0.1mg/ml) plasmid was added and mixed gently by flicking the tube. After leaving on ice for 30 minutes the cells were heat-shocked for 45 seconds in a water bath at 42°C and immediately returned to ice for 2 minutes. The cells were suspended in 400μl Luria Broth[3] (LB), a medium rich in nu-

[3]LB is also known as Lysogeny Broth and Luria Bertiani media

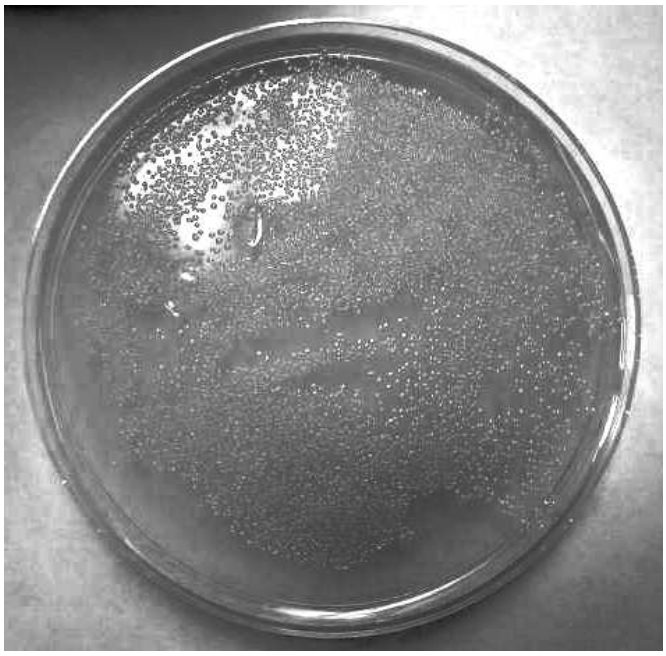


Figure 3.2: An LB agar plate after an overnight incubation. Each small dot is a colony grown from an individual cell. Single colonies can be carefully removed and used to grow a pure culture.

trients required for the growth of bacteria, and grown for an hour on a shaker at 37°C. The solution was spun in a Heraeus Biofuge Stratos centrifuge at 12000g , 4°C for 1 minute to pellet the cells, which were then resuspended in 50μl LB. Five 10μl samples were plated on five LB agar plates supplemented with 100μg/ml carbenicillin and 34μg/ml chloramphenicol and incubated overnight at 37°C to produce single colonies.

Agar is a stiff, polysaccharide gel that can be combined with LB media to grow colonies on a surface. Specific compounds like antibiotics may also be added to apply a selection pressure, which biases growth towards the microorganism of interest. LB agar is usually prepared in a shallow dish called a Petri dish. A drop of the cell culture is spread thinly, or "streaked", across the surface of the LB agar and individual colonies grow from the dispersed cells (figure 3.2). This technique is referred to as “plating”, where the LB agar in Petri dish is a “plate”. A single, isolated colony is then selected and used to grow a large volume of pure culture.

3.2.2 Transcription

Traditionally LB media is used to grow the culture. A single colony is carefully removed from the plate and used to inoculate 5ml LB supplemented with antibiotics (100μg/ml carbenicillin and 34μg/ml chloramphenicol). This is incubated overnight at 37°C with shaking. The process is repeated for 100ml LB using samples from the overnight cultures and then again for 500ml LB, both solutions containing the usual antibiotics. The latter culture is grown in 2L baffled flasks on a shaker until the O.D.^{600} (Optical Density at 600nm) ≈ 0.6. Optical density is a measure of the growth of the culture and is defined as the absorbance per unit length, which would be the cuvette width for a UV-spectrometer. Once the required OD has been reached transcription can be induced by adding IPTG (isopropyl-β-D-thiogalactoside) at a concentration of 1mM and incubating and shaking for a further 2 hours. Finally, the cells are harvested in a centrifuge and the pellet can be frozen at -20°C for use at a later date.

Induction can be thought of as switching on the transcription machinery of the cell. A section of the transformed DNA called the operon is used to ensure that OCR is only produced when IPTG is present. IPTG can therefore be added when the culture has reached its peak growth to maximise the amount of OCR produced. An operon is a set of key nucleotide sequences (a promoter, an operator and the genes for the required proteins) that control the production of messenger RNA (mRNA). The promoter sequence is recognised by RNA polymerase, which associates with it and carries out transcription. The operator acts as a switch for the transcription by interacting with a repressor protein. This binds to the operator sequence and in doing so covers up part of the promoter, blocking RNA polymerase and preventing transcription. IPTG binds to the repressor protein and inactivates it. The repressor is prevented from binding to the operator and RNA polymerase is free to transcribe the genes. Transcription is said to be induced by IPTG. In standard *E.coli* cells with untransformed DNA, induction is regulated by lactose.

In the latter parts of the work, I used a new media for culture growth called Overnight Express produced by Novagen UK. It automatically induces protein expression once the cells reach stationary phase, when the nutrients in the medium begin to run out and the bacterial growth rate slows. Five single colonies were grown in five separate 5ml LB solutions overnight as described above. To test the cultures for the best induction a 500μl sample was taken from each. IPTG was added to a concentration of 1mM and

put on a mixer for 2 hours at 37°C. The sample was spun in a centrifuge at 30000g, 4°C for 10 minutes and the cell pellet resuspended in 500μl of 20mM Tris-HCl, pH 8 . After sonicating for 30 seconds (at $\sim$10mW) and spinning at 30000g for 10 minutes the supernatant and 5ml culture were loaded onto a 15% SDS polyacrylamide gel. Electrophoresis is used to separate all the elements of the induced and uninduced cultures (see section 3.1.1). A marker ladder and a control sample of pure ocr were used as references. A molecular weight marker contains different proteins with a range of known sizes. They appear on an electrophoresis gel as a ladder that can be used to approximate the size of unknown proteins. Samples 3 and 4 show the best induction (strongest band on the gel), as shown in figure 3.3. 0.5ml of bottle 3 was used to inoculate 500ml of Overnight Express in a 2L baffled flask. The flasks were placed on a shaker at 37°C and incubated overnight, during which protein expression was automatically induced. The cultures were then spun at 30000g for 10 minutes to harvest the cells. An 8L culture gives around 20g of cell pellet. The pellet is ready for the protein purification stage or it can be stored at -20°C. All bottles, flasks and supernatant waste should be disinfected with Virkon.

3.3 Protein purification

3.3.1 Cell lysis

To release the soluble protein into solution the cells need to be broken open. This can be achieved efficiently using a combination of lysozyme and ultrasound. Cells experience a high osmotic pressure that causes them to expand. This internal pressure is caused by a lower concentration water and a higher concentration of biological molecules (proteins, DNA, ribosomes etc.) inside the cell than in the surrounding environment. The expansion is contained by a tough cell wall that supports the cell's delicate membrane. In bacteria the cell wall is a stiff layer of long polysaccharide chains, cross-linked by short amino acid strands into a lattice. Without this protective barrier to brace against, the cell would burst under its own internal pressure. Lysozyme exploits this weakness to destroy bacteria. It cleaves the polysaccharide chains by hydrolyzing the glycosidic bond between two adjacent sugar molecules. The cell wall is gradually dissolved, destroying its structural integrity, until it ruptures and the cell bursts open [97, 98].

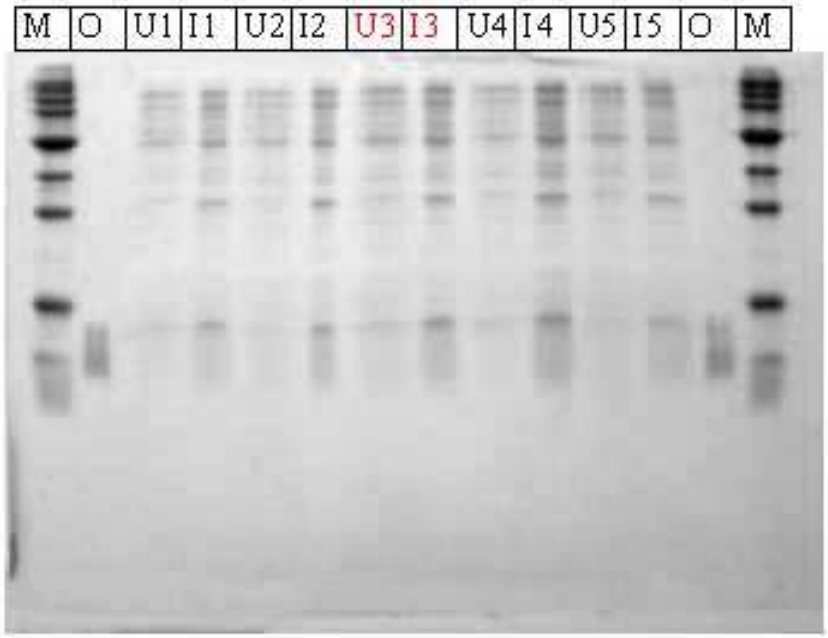


Figure 3.3: The electrophoresis gel from testing the five 5ml LB cultures. The 14 vertical columns represent the uninduced and induced cultures along with two reference sets. As the proteins in each sample (highlighted by horizontal bands) move down through the gel they are separated by size (the smallest at the bottom). While all the induced samples (I1-5) show strong OCR bands, I3 was chosen to inoculate Overnight Express media.

Key: Column M contains the molecular weight marker ladder (10-255kDa, figure 3.1), O contains a pure sample of ocr, U1 is the uninduced culture and I1 the induced culture from sample 1 (similarly for the other numbers).

A lot of DNA is released during cell lysis, resulting in a highly viscous solution. A short amount of sonication will break up the DNA, reducing the viscosity. It also has the added effect of lysing any remaining cells that may have escaped the lysozyme. Sonication should be done in short pulses as ultrasound can raise the temperature quickly, especially in the vicinity of the probe. The tube containing the cell suspension should also be surrounded by ice to keep it cool. After sonication the unwanted, heavy products like DNA, insoluble material and unlysed cells are removed by spinning in a centrifuge.

The frozen cell pellet was thawed on ice for 15 minutes and resuspended in buffer A (20mM Tris-HCl, 0.3M NH_4Cl, pH 8) using 5ml buffer per gram pellet. Lysozyme was added to a concentration of 200μg/ml (easier to use a stock solution of 2mg/ml than powder). PMSF[4] and benzamidine were added to 20μM and 10μM respectively after 10 minutes to inhibit proteases released during lysis. The cell suspension was stirred slowly using a magnetic flea for 30 minutes at 37°C. When the solution became very viscous it was put on ice and sonicated for 2 minutes, alternating between 30 seconds on and 30 seconds off. Spinning at 20000g, 4°C for 1 hour separated the cell debris and insoluble waste from the protein solution.

3.3.2 Precipitation cuts

This is a simple and quick way of removing a significant number of unwanted proteins based on screening in high salt conditions (see section 2.5). On average most proteins precipitate before 2M ammonium sulphate, with the exception of ocr. Two cuts were made. One at 1.2M (30% saturation) and one at 1.8M (45% saturation) ammonium sulphate. The salt was dissolved slowly in the supernatant and stirred for 10-30 minutes. To remove the precipitated proteins the sample was spun at 20000g, 4°C, for 10 minutes. This process was then repeated for the next cut. Any salt remaining in solution was removed by dialysis against buffer A. 10000 MWCO[5] SnakeSkin dialysis tubing from Pierce was used in 5L of buffer, for 50ml samples. Dialysis was performed at 4°C for 3 hours, replacing the buffer every hour.

[4] phenylmethylsulfonyl fluoride

[5] Molecular Weight Cut Off. A value half the weight of the protein is recommended to avoid losses.

3.3.3 Ion exchange chromatography

Chromatography is the method for separating solutes according to the amount of time it takes each species to travel through an adsorbing material. A vertical glass column is filled with the material and the sample filters down through it. The extent by which the solutes are adsorbed within the column dictates their speed of descent, resulting in the different species being separated in time.

Ion exchange chromatography is ideal for separating charged compounds like proteins. The material consists of a matrix of charged groups. The oppositely charged protein of interest can reversibly bind to these ion exchange groups. Initially the column is equilibrated in a state that allows the desired proteins to bind to it. The sample is applied to the top of the column and the solutes of opposite charge are reversibly bound while the contaminants pass through freely. The ionic strength of the sample buffer is gradually increased to progressively remove the bound substances. This is called a gradient and is usually achieved by increasing the salt concentration. The most weakly bound solutes are washed out first. The solution leaving the column is collected in fixed volumes called fractions. These are analysed using a combination of UV spectroscopy (see 3.1.2) and gel electrophoresis (see 3.1.1) to reveal the location of the required protein.

Ion-exchange chromatography with a DEAE[6]-cellulose column was used to separate ocr from any proteins remaining in solution. This is a very effective method as the highly acidic ocr protein has a high affinity for the positively charged DEAE-sepharose. Initially the column was cleaned with the high salt buffer B (20mM Tris-HCl, 1.0M NH_4Cl, pH 8) and equilibrated with the low salt buffer A (see 3.3.1). Two column volumes of buffer were used at each stage with a flow rate of 1ml/min. Rates higher than 1ml/min are not recommended as they reduce the protein separation and may compact the column. The protein sample was loaded onto the column and the fluid leaving the column (*flow through*) collected for testing later. The column is then washed with 4 column volumes of buffer A to remove the unbound substances. Once again the displaced fluid (*wash*) is collected. The protein is then eluted with a 500ml gradient from low to high salt buffer. One hundred 5ml fractions are collected. The column is washed with buffer B and stored in 20% ethanol.

Figure 3.4 shows the data collected during the chromatography. A clear peak in

[6]diethylaminoethyl

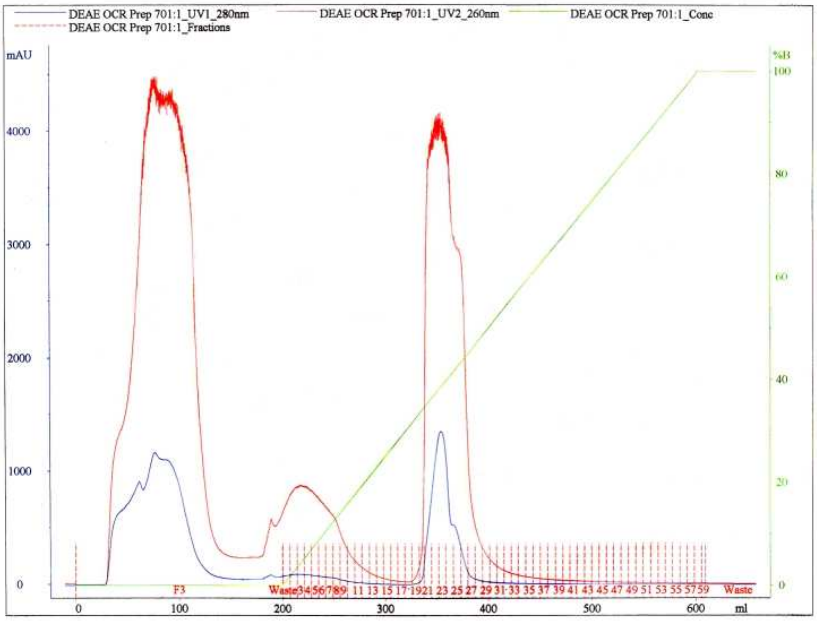


Figure 3.4: Data recorded during the DEAE chromatography. UV absorption by proteins at 280nm (blue, lower line) and DNA at 260nm (red, upper line) is shown alongside the green, straight line depicting the change in buffers during the gradient and the red markers showing the fraction numbers. Ocr is detected from fraction 20 and eluted over fractions 23-29. There is a large amount of DNA present.

the UV absorbance at 280nm can be seen during the gradient around fractions 20-40, corresponding to ocr. There is also a lot of DNA remaining, as revealed by the UV absorbance at 260nm. Fractions 3-40 were tested on an electrophoresis gel (figure 3.3) along with the cell lysis, pellet and supernatant from the previous stage and the column flow through and wash. Fractions 24 - 27 were chosen as they contain the most ocr with as few contaminants as possible. The selected fractions were pooled and a UV absorption spectrum taken (figure 3.6). A significant amount of DNA is present as the absorption at 260nm is far greater than at 280nm ($A_{280} : A_{260} = 0.588$). This will need to be removed before the protein can be used in further experiments.

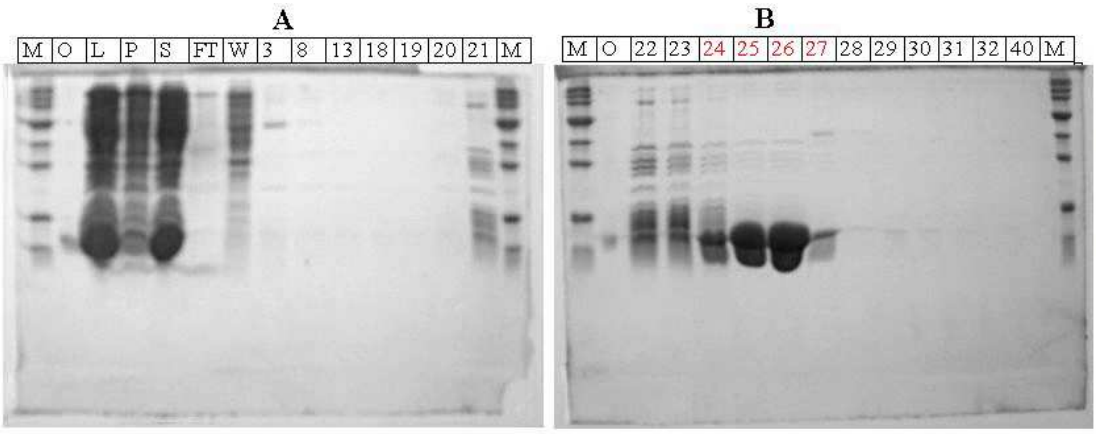


Figure 3.5: Electrophoresis gels showing the result of a DEAE column purification. Each vertical column represents a sample from the preparation stages or selected DEAE fractions. The horizontal bands represent the proteins in each sample, separated by size. Fractions 24 - 27 were chosen as they contain the most ocr with as few contaminants as possible. The samples appear overloaded at the peak fractions but this reveals the amount of minor proteins present.

Key: (left to right) Gel A - marker (M), ocr (O), cell lysis before centrifuging (L), pellet (P) and supernatant (S) after centrifuging, flow through (FT) and wash (W) from column, fractions 3 - 21 and another marker; Gel B - marker (M), ocr (O), fractions 22 - 40 and a final marker.

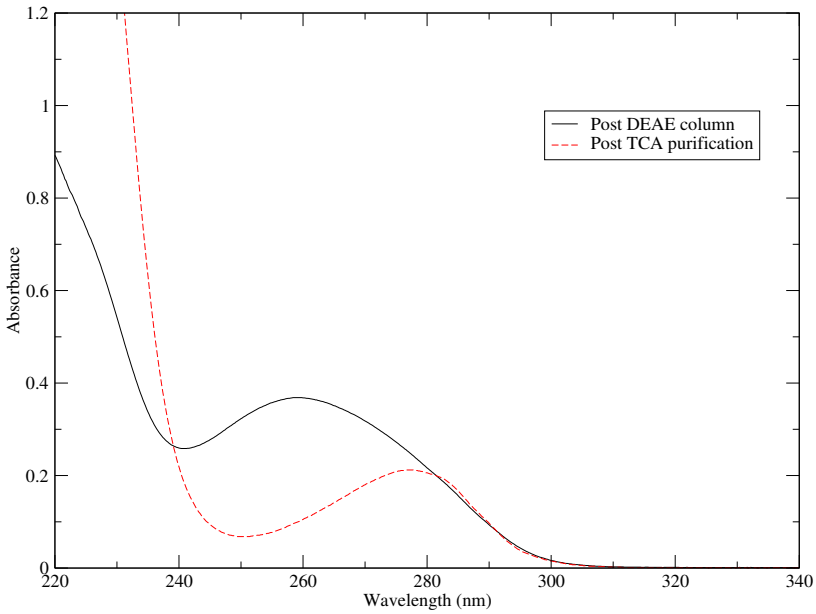


Figure 3.6: UV spectrum of the pooled, chosen fractions from the DEAE column and the final solution after TCA/ethanol purification. A 10mm cuvette was used containing a 30$\times$diluted sample. The peak of the DEAE trace is clearly centered on 260nm with a ratio of $A_{280} : A_{260} = 0.588$, revealing that a large amount of DNA is present. The situation is much improved after TCA precipitation with the peak shifting to 280nm and a ratio of $A_{280} : A_{260} = 1.96$. The final protein concentration was 0.2mM = 5.4mg/ml.

3.3.4 Removing DNA

3.3.4.1 More chromatography

Hydrophobic chromatography works in the same way as ion exchange but employs hydrophobic interactions instead of electrostatic ones. The hydrophobic patches on proteins prefer to be in contact with other hydrophobic surfaces, for example phenyl sepharose. The interaction is strongest at high salt concentrations (e.g. 2.4M $(NH_4)_2SO_4$), which are used to equilibrate the column. A gradient to 0.0M separates the solutes according to their hydrophobicity.

Gel filtration is also known as size exclusion chromatography because it separates solutes based on their size. The column material is a gel containing pores, which have a controlled range of sizes. For example a superdex 75 resin has pores between 24-44μm in diameter. As the solution filters down the column solutes diffuse into the pores to varying extents. The largest solutes cannot diffuse in at all and remain in the solution outside. These travel fastest through the column and are the first to leave, while the smaller solutes take longer to exit.

Both techniques were attempted but generally give poor resolution and unsatisfactory separation of DNA from ocr. However, ocr is soluble in ethanol and this leads to a simpler purification method.

3.3.4.2 TCA precipitation

Trichloroacetic acid precipitation with ocr was pioneered by Mark and Studier and has proved to be the most effective method for separating DNA from ocr [60, 86]. The pooled fractions from the DEAE column were precipitated with an equal volume of 10% trichloroacetic acid (TCA). After 10-20 minutes the precipitate was collected in a centrifuge (28000g for 30min) and redissolved in 95% ethanol. The contaminants (DNA and rogue proteins) were insoluble in ethanol and were easily removed using a centrifuge (28000g for 30min). This cycle is repeated and the ocr recovered by precipitating with TCA one final time and dissolving the precipitate in a 20mM Tris-HCl, pH 8, buffer. Any remaining salts, TCA or ethanol can be removed by dialysis against the Tris-HCl buffer. The final solution is concentrated using VIVA spin-column filters with a MWCO of 5000 in a centrifuge. Although ocr is a highly stable protein, as always this procedure was performed at 4°C. A UV spectrum for the final solution

was taken (figure 3.6) and shows a significant improvement. The absorption peak has shifted to 280nm and the absorption at 260nm due to DNA is greatly reduced ($A_{280} : A_{260} = 1.96$). The final protein concentration was 0.2mM = 5.4mg/ml and the whole process yielded around 146mg ocr in 27ml.

The TCA and ethanol treatment may not seem very gentle but Mark and Studier have shown it to have little effect on ocr [60]. Its specific ability to inhibit the EcoKI type I R-M enzyme activity was unaltered. Ocr also remained a dimer in solution. An equal volume of glycerol was added to the purified protein (50% v/v glycerol/protein solution), which was stored at -20°C.

3.3.5 A final word

The expression and purification of ocr is both time and labour intensive. Particularly before the improvements discussed in this chapter were implemented. Initially a week and a half was required to produce a mere 10mg of protein. One rheology experiment required at least 1mg ocr, while circular dichroism required 0.3mg and atomic force microscopy 0.1mg, not to mention ocr used in perfecting the gelling procedure, performing preliminary observations and other more minor experiments. Since a large range of conditions were probed, and each experiment repeated a number of times for reproducibility, hundreds of miligrams of ocr were required for this work. After improving the efficiency and yield of the expression and purification it was possible to produce 100mg in just under a week. However, this procedure still had to be performed once every couple of months as stocks of protein were quickly used up in experiments.

Chapter 4

Experimental methods

The ocr gel was discovered accidentally by David Dryden and Shane Sturrock, during a routine precipitation check. Solid ammonium sulphate was added to a solution of ocr, which remained transparent but simultaneously became viscous. To investigate this unusual behaviour better quality samples were required, free from the bubbles produced by disolving salt granules. I tried and refined a number of techniques to produce a transparent gel, free of contaminants. These are described in the results section 5.1.

Once a pure ocr gel was obtained a number of different methods were required to characterise it. The physical nature of the gel, in particular its "gel" status, was investigated by rheology (section 4.1). However, this only provided limited insight into its underlying structure. The spectroscopic technique circular dichroism (section 4.2) and atomic force microscopy (section 4.3) offered less intrusive methods to study the gel. They revealed the formation of long, helical fibrils that formed a space-spanning network at the gel point. The narrow range of conditions under which the gel forms were also discovered using these techniques. Three other experiments were attempted: particle tracking, fluorescence anisotropy and fluorescence microscopy. These refined the insights obtained by the other methods, and are discussed in chapter 6.

4.1 Rheology

4.1.1 Introduction

The behaviour of a system away from thermodynamic equilibrium is often where its most interesting aspects lie. The response of a material to an external force also has important industrial consequences. The squeezing of toothpaste, the consistency of food and the spreading of paint all illustrate the need for an understanding of how stresses and strains affect a material. The investigative tool for these processes is rheology: the study of deformation and flow under external mechanical forces [99, 72, 100].

The two extremes of rheological behaviour are characterised by the basic notions of liquids, which flow, and solids, which do not flow. When an ideal solid experiences an external mechanical force it instantaneously deforms by a finite amount. The structure of the material gives it an "internal resistance", which opposes further distortion and stores the energy that caused the initial deformation. This energy is recovered when the external force is removed. The deformation spontaneously reverses and the solid relaxes to its original form. In contrast, the same force applied to a simple fluid causes a continuous deformation because a fluid's internal resistance relaxes almost instantaneously. The fluid flows while the external force is applied and stops when the force is removed. However, the deformation is not reversed and the fluid does not recover its original form; energy is dissipated during the flow as heat. In this case the internal resistance no longer limits the deformation itself but the *rate* of deformation.

These two types of limiting behaviour are called elastic and viscous behaviour. Between them lies a range of systems whose behaviour depends on the time scale over which the force is applied. These systems exhibit viscoelastic behaviour. A common example is a ball of silicone rubber, known as "silly putty". It displays elastic behaviour by bouncing off a wall, where the contact time is a few milliseconds, and viscous behaviour by deforming when squeezed or pulled for a few seconds. Even glass, traditionally viewed as a solid, will flow over hundreds of years. This causes the characteristic melted appearance at the foot of tall church windows, where the pane is thicker than at the top. Indeed, the motto of The Society of Rheology is "everything flows", with the unspoken qualifier "given long enough". "Long enough" for a given material depends on both the characteristic relaxation time of its internal resistance, $t_{relaxation}$, and the time taken to make an observation, $t_{observation}$. The time it takes for a

system to relax ranges from $t_{relaxation} \to 0$ for a simple viscous fluid (water is typically 10^{-12}s) to $t_{relaxation} = \infty$ for an ideal elastic solid. The ratio of these two time scales defines the Deborah number [101]:

$$De = \frac{t_{relaxation}}{t_{observation}} \tag{4.1}$$

It provides a general description for the behaviour of a certain material over the time scales involved. A high Deborah number corresponds to predominantly elastic, solid-like behaviour, while a low De indicates viscous, fluid-like behaviour. A material with a low relaxation time may therefore appear solid-like when observed with a very fast deformation process. For Deborah numbers around unity a wide range of viscoelastic behaviours are observed.

The Deborah number reveals how important rheology is for the study of colloidal systems. In this case the internal resistance relaxes due to Brownian motion of the dispersed particle. They follow a random walk where the average mean-squared displacement of a spherical particle over a time Δt is given by $< \Delta r^2 > \sim D\Delta t$, for a diffusion coefficient D. The relaxation time can be said to equal the time it takes a particle to diffuse its own radius, r: $t_{relaxation} \sim r^2/D$. This can range from milliseconds for dilute particle suspensions to hours for concentrated ones. Since these are comparable to the usual time-scales for rheological observations, $De \sim 1$ and we can expect a variety of viscoelastic phenomenon from colloidal systems.

Viscoelastic materials combine characteristics from elastic solids that store energy and viscous fluids that dissipate energy. They can flow under a constant external force but some mechanical energy is conserved and can be recovered once the force is removed. The underlying structure causing such behaviour can be varied and complex. To develop the rheological theory we return to an ideal material and basic deformation, that of a Newtonian fluid under simple shear.

4.1.2 Newtonian fluid under simple shear

Figure 4.1 illustrates the concept of simple shear. A tangential force is applied to the top surface of a cuboid, while the bottom remains fixed, deforming it into a rhomboid. If instead, we consider a pack of playing cards, a simple shear results in a tilted stack. The tangential force, F, acting on the top card, of surface area A, produces a shear stress, σ, defined as:

$$\sigma = \frac{F}{A} \tag{4.2}$$

In the tilted stack, each card has a thickness δy and each is shifted a distance δx beyond its lower neighbour. The shear deformation has caused a shear strain, γ, defined as:

$$\gamma = \frac{\delta x}{\delta y} \tag{4.3}$$

In an ideal solid, which resists deformation, the shear stress results in a fixed δx and the shear strain is constant. A larger force, or one that acts over a smaller area would result in a proportionally larger deformation. In the limit $\gamma \to 0$, shear stress is therefore proportional to shear strain with a constant of proportionality called the rigidity, or shear, modulus of the solid:

$$\sigma = G\gamma \tag{4.4}$$

In a viscous fluid a constant tangential force, or shear stress, would induce a steady shear flow. The fluid can still be treated as a stack of very thin layers but now each moves at a constant (yet different) velocity with the tangential separation between layers increasing with time, $\delta\dot{x} = \frac{\delta x}{\delta t}$. We can envisage a fluid trapped between two parallel plates, separated by a vertical distance Y. The lower plate is fixed and the upper plate travels with a constant velocity, V, due to the applied shear stress, σ. Assuming a condition called "no-slip boundary conditions", the layer of fluid in contact with the lower plate is at rest, while the layer in contact with the upper plate travels at velocity V. A velocity gradient exists throughout the fluid that depends on its composition. In the simplest case, the imposed shear stress results in a constant shear rate, $\dot{\gamma}$, given by:

$$\dot{\gamma} = \frac{\delta\dot{x}}{\delta y} = \frac{V}{Y} \tag{4.5}$$

The shear stress in a Newtonian fluid is directly proportional to the shear strain *rate* (as opposed to the shear strain in the case of a solid, equation 4.4):

$$\sigma = \eta\dot{\gamma} \tag{4.6}$$

where η is the viscosity, a measure of the fluid's resistance to flow. The Newtonian fluid model is a good description for many low molecular weight liquids, including the

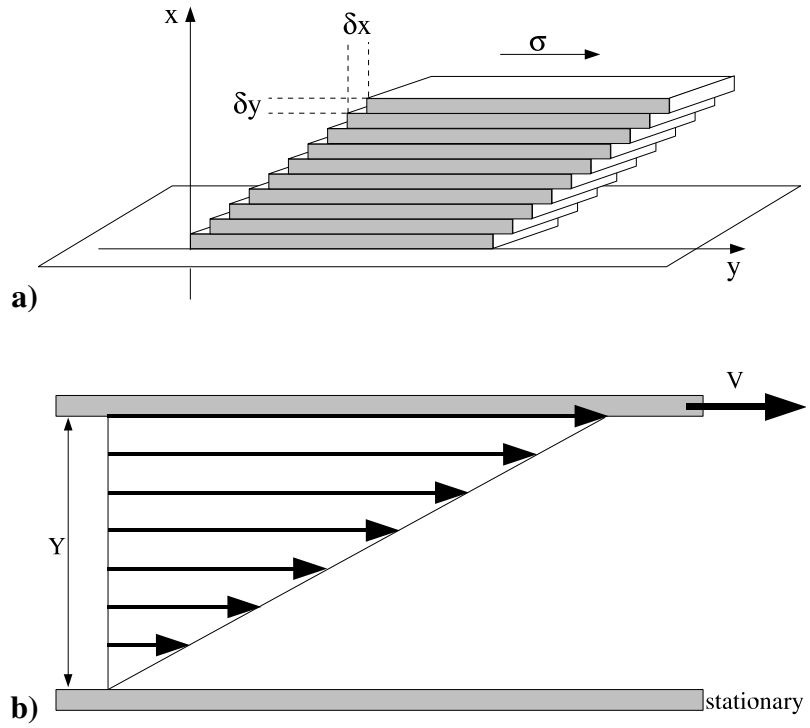


Figure 4.1: **a)** Illustration of shear showing a cuboid being deformed into a rhombus by a shear force imposed on the top surface. The cuboid can be thought of as consisting of many thin layers, each of thickness δy and displaced from the lower layer by δx. **b)** Illustration of a viscous fluid trapped between a stationary plate and a plate that moves at velocity V. A velocity gradient exists in the fluid, which for a Newtonian fluid is constant. The shear rate is then given by $\dot{\gamma} = \frac{\delta \dot{x}}{\delta y} = \frac{V}{Y}$.

familiar examples of water ($\eta = 10^{-3}$Pa.s), olive oil ($\eta = 10^{-1}$Pa.s) and golden syrup ($\eta = 10^{2}$Pa.s) [99]. For most fluids, however, the viscosity is not simply a constant of proportionality but a function of shear rate, $\eta(\dot{\gamma})$. These materials are called complex fluids and exhibit non-Newtonian flow behaviour.

4.1.3 Non-Newtonian behaviour

Steady shear flow is the simplest method for probing an unknown material. If a constant shear stress is applied to the sample the resulting steady state shear rate can be measured. This is done over a range of stresses to obtain the material's flow curve,

a graph of shear stress verses shear rate[1], $\sigma(\dot{\gamma})$. Since the hallmark of Newtonian behaviour is a shear viscosity that does not vary with shear rate a Newtonian fluid flow curve will be a straight line whose constant gradient is the viscosity, $\eta = \frac{\sigma}{\dot{\gamma}}$. This is shown in figure 4.2a)(i), along with three other characteristic flow curves. Figure 4.2b) shows the corresponding (apparent) viscosities for each flow curve. If the viscosity during steady shear flow increases with increasing shear rate the material is said to exhibit shear thickening behaviour (ii). Conversely, shear thinning behaviour occurs if the viscosity decreases with increasing shear rate (iii). These situations are often described by the power-law model [102, 10, 103]:

$$\sigma = m\dot{\gamma}^n \tag{4.7}$$

where m and n are constants that are obtained by fitting to the data. The model describes shear thinning for $n > 1$ and shear thickening for $0 < n < 1$. If $n = 1$ the power-law model describes a Newtonian fluid where $m = \eta$, as in equation 4.6. The final curve (iv) in figure 4.2 represents a material that behaves under simple shear like an elastic solid ($\eta \to \infty$) up to a critical stress, called the yield stress σ_y, whereupon it flows with a viscosity called the plastic viscosity, $\eta \to \eta_{pl}$. This behaviour is described by the Bingham model:

$$\sigma = \sigma_y + \eta_{pl}\dot{\gamma} \tag{4.8}$$

Measurements of the yield stress can be improved by using the creep method described in section 4.1.4. Non-Newtonian behaviour is frequently observed in colloidal suspensions and concentrated polymer solutions. Shear thinning under steady shear flow is a common property for many materials, particularly among polymeric liquids. Everyday examples include toothpaste and mayonnaise, which may easily be squeezed or spread but will not flow under small forces like gravity. Structurally it represents an decrease in isotropy, rearranging particles to lower the flow resistance in the direction of shear [99]. At low shear rates particles in a concentrated suspension collide during flow. Brownian motion dominates and the fluid has a high viscosity. As the shear rate increases an orientation of particles is induced that Brownian motion cannot overcome. Particles move more freely and the viscosity decreases. This can also be seen in so-

[1] Flow curves are traditionally plotted as stress verses shear rate, regardless which variable is used as the control

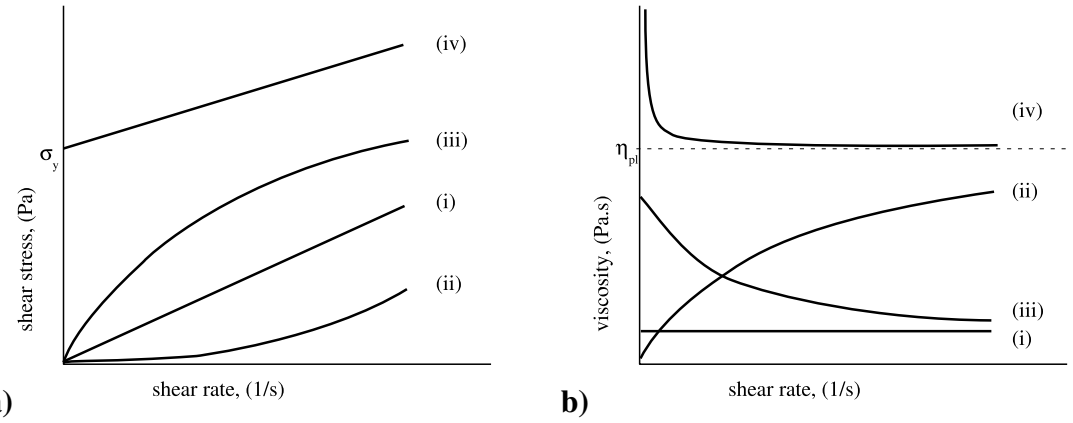


Figure 4.2: **a)** Schematics of characteristic steady shear flow curves and **b)** the corresponding variation in viscosity with shear rate; (i) Newtonian, (ii) shear thickening, (iii) shear thinning and (iv) Bingham with a yield stress σ_y and plastic viscosity η_{pl}.

lutions of polymers and fibrils where they align along the direction of flow. Materials may also demonstrate a combination of these distinct behaviours over different shear rate regions. For example, concentrated hard sphere suspensions have been observed to show shear thinning at low shear rates where the particles can move around each other through Brownian motion, assisted by a gentle shear force. As the shear rate is increased they have no time to relax, the particles cluster together and the material shear thickens [104].

4.1.4 Creep

Instead of varying the shear stress imposed on a material, as in the case for shear flow, creep involves fixing the shear stress and measuring how the strain develops over time. Ideally the step increase in stress should occur instantaneously. For mechanical reasons there is always a small delay in reaching the target shear stress but this should be significantly less than the material's internal relaxation time. The stress is typically held constant for minutes or hours. When the stress is stopped suddenly the sample will try to recover its original shape. Whereas an ideal elastic material will not flow but will completely recover its initial shape and an ideal viscous material flows steadily but has no recovery, a viscoelastic material will flow and partially recover. Some permanent deformation remains, depending on the sample's structure and the force to which it is subjected [103]. These effects are summarised schematically in figure 4.3.

By performing the creep experiment for a range of shear stresses the minimum value required for flow (the yield stress) can be calculated. The data from each experiment is plotted on a strain vs. time graph and a power law of the form $y = Ax^B$ fitted to those data points within the steady state region (as shown in figure 4.4). The power coefficients, B, for each experiment are then plotted against their corresponding stress. B is related to the internal resistance of the material. When it is low the sample is resisting the deformation at that shear stress. The material may not be able to support higher stresses and the point at which B starts to increase rapidly is called the yield stress. At $B \sim 1$ a steady state Newtonian flow has been created in the fluid.

Yield stress can be thought of as the transition from elastic to viscous behaviour and is dependent on the internal network structure. It is commonly regarded as occurring at a single point but there is some debate as to its true definition or even existence [105]. In some circumstances, particularly if a limited range of shear rates are probed, shear thinning may appear similar to yield stress, if the rearranging of particles occurs rapidly above a specific stress. However, it is still a useful concept for investigating materials whose constituent parts interact to form a continuous network. When stresses high enough to overcome its elasticity are applied the structure breaks down and the material flows. The point (or spread of stresses) at which this occurs depends on the concentration and extent of particle aggregation or polymer entanglement.

During creep experiments the strain generally increases slowly initially (the creep region). After some time the accumulated strain is too large for the structure to bear elastically and the material starts to flow, resulting in a sudden increase in strain rate. If a stable, viscous flow is established the strain rate becomes constant. The larger the imposed stress the sooner the material yields. The elastic properties of the material may be recovered when the applied stress is removed. The extent to which this happens depends on the degree of failure in the structure caused by the stress.

4.1.5 Oscillatory flow

Creep experiments are known as static methods because they impose a single step change and measure the subsequent response over time. Dynamic methods, on the other hand, apply a harmonically varying stress. They have the advantage of being able to probe a large range of relaxation times. If the frequency range available is from 10^{-3} to $10^{3}\mathrm{s}^{-1}$, then time scales from about 10^{3} to 10^{-3}s can be covered.

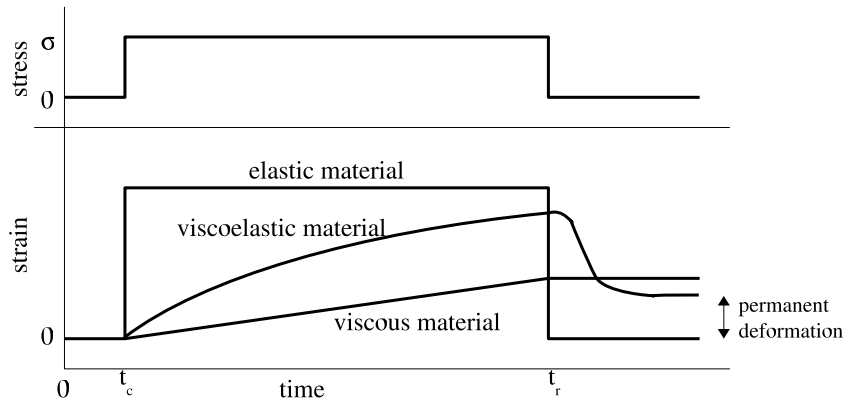


Figure 4.3: Creep and recovery curves. A fixed stress is applied at time t_c and stopped at time t_r. An ideal elastic material does not flow and recovers completely when the stress is removed. An ideal viscous material experiences a steady flow but no recovery. An ideal viscoelastic material also flows but when the stress is removed it partially recovers its original form.

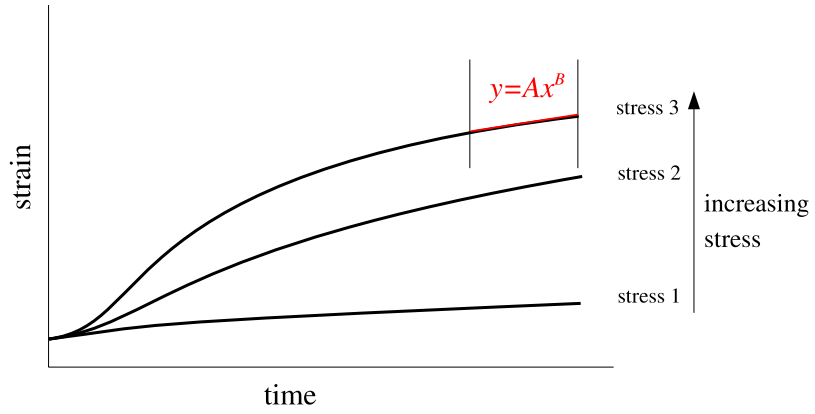


Figure 4.4: To find the yield stress of a material perform a number of creep experiments on a sample, with increasing stress. Fit a power law of the form $y = Ax^B$ to the steady state region of each experiment (highlighted in the figure), then plot a graph of B against stress. The stress at which B starts to increase rapidly is called the yield stress.

When using oscillatory probes, experiments must be performed in the linear viscoelastic region of the material so that its structure remains intact. This regime is usually accessed by applying a range of oscillatory stresses and focusing on the region where the moduli are independent of the applied stress (see figure 5.6). Working within the linear region, the frequencies of the stress oscillations are varied to obtain the storage and loss moduli. A material's shear modulus is given by the ratio of stress and strain. For oscillatory experiments the applied stress and resulting strain vary sinusoidally as follows:

$$\begin{aligned} \sigma &= \sigma_\circ \cos \omega t \\ \gamma &= \gamma_\circ \cos(\omega t - \delta) \end{aligned} \tag{4.9}$$

The strain developed in a material due to an applied oscillatory stress may be out of phase with the stress by a phase angle δ. This may vary between the limits $\delta = \frac{\pi}{2}$ for a Newtonian fluid and $\delta = 0$ for a solid. Equation 4.9 can be expressed in terms of complex variables where $\gamma_\circ e^{(-j\delta)}$ is called the complex strain amplitude. The complex stress amplitude remains $\sigma_\circ$ and the ratio of these defines the complex shear modulus G^*:

$$G^* = (\frac{\sigma_\circ}{\gamma_\circ} \cos \delta) + (\frac{\sigma_\circ}{\gamma_\circ} i \sin \delta) \equiv G' + iG'' \tag{4.10}$$

The storage modulus G' is the in-phase, elastic response of the system while the loss modulus G'' is the out-of-phase dissipative response [102, 10]. We can distinguish between the general behaviour of a viscoelastic material based on the frequency dependence of these moduli. Solid-like behaviour is observed if G' remains finite as $\omega \to 0$, and $G' > G''$, resulting in a high stability against mechanical influences. When $G' < G''$ the material exhibits liquid-like behaviour under oscillatory shear and is able to undergo large deformations and flow. An example of this transition, obtained by Sood *et al.,* can be seen in figure 4.5, which describes the rheology of a viscoelastic network of elongated micelles formed from the surfactant CTAT (Cetyltrimethylammonium Tosilate) [106]. The material has mainly elastic properties at low frequencies, where $G' > G''$, but above a critical oscillatory strain rate (0.45rad/s) the behaviour becomes more like a viscous fluid, where $G' < G''$.

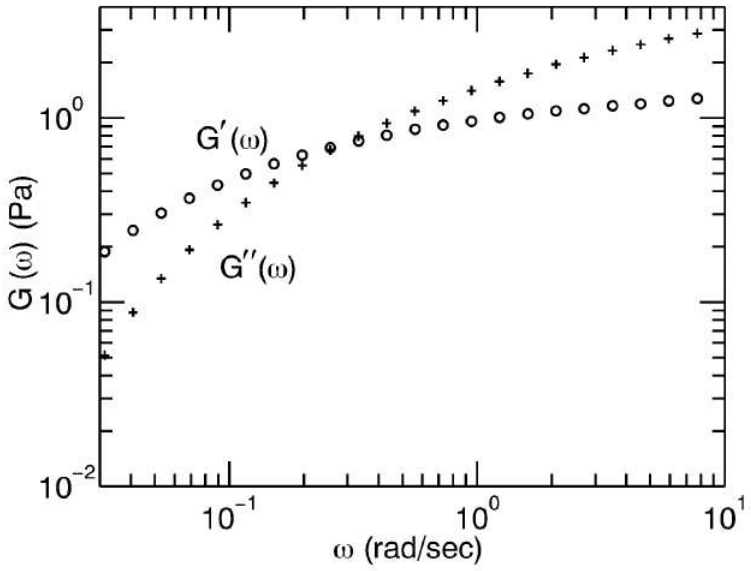


Figure 4.5: An example of storage and loss moduli for a viscoelastic network, showing a crossover between predominantly elastic behaviour ($G' > G''$) to viscous fluid ($G' < G''$). The system was investigated by Sood *et.al.* and involved a solution of entangled, elongated micelles at 25°C. Diagram reproduced from [106].

4.1.6 Apparatus

Rheology investigates the relationship between stress and strain (deformation) or strain rate (flow). It is conventionally performed using a rheometer, which applies a precisely controlled, mechanical shear force to the material and measures its response. The material is usually sandwiched between a fixed surface and a surface that rotates, shearing the sample (figure 4.6a). The relationship between stress and strain or strain rate depends on the geometry of the surfaces and on measurements of the torque, which is proportional to the stress, the relative angular displacement (proportional to the strain) and the angular displacement per second (proportional to the strain rate) of the rotating surface.

To measure rheological properties, a rheometer may directly control either the shear stress applied (a stress-controlled rheometer), or the strain rate applied (a strain-controlled rheometer). Regardless of the rheometer type, a torque imposed by a motor is required to rotate the surface; the manner in which the torque is controlled determines the properties of the rheometer. With strain-controlled rheometers, a motor controls the displacement and velocity of the rotating surface to achieve the required

strain or strain rate, while the torque is measured through a separate transducer. In contrast, with stress-controlled rheometers a constant torque motor is used to generate a torque that is independent of the rotational speed. To measure the strain or strain rate induced in the material an optical encoder or radial position transducers may be directly attached to the rotating device. Stress-controlled rheometers provide accurate measurements at low stresses, during long-time fixed-stress (creep) experiments and when the strain or strain rate varies rapidly over many orders of magnitude [11, 107].

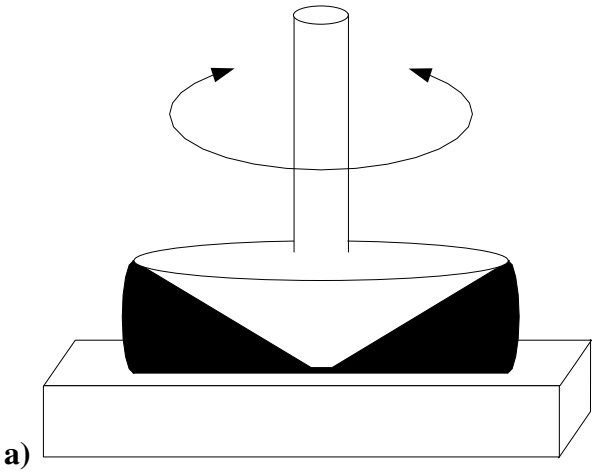
In this work a stress-controlled *AR2000* rheometer from *TA Instruments* was used (figure 4.6b). This was also able to perform "strain controlled" experiments, like steady shear flow, by a sensitive feedback loop that continuously measures the strain induced in the material and alters the applied stress to maintain a constant strain.

Two horizontal, parallel plates are often used to contain the sample and shear it, or alternatively a Couette cell, which consists of two concentric cylinders, the outer one stationary, the inner rotating. The most common geometry employed though is a cone-and-plate system as it gives a constant shear rate over the whole sample [108]. This is because further out from the centre of the cone the sample moves faster but since the gap between the cone and stationary plate has increased the ratio of velocity to height remains the same. By contrast for parallel plates and Couette cells, the shear rate increases with the velocity along the radial direction.

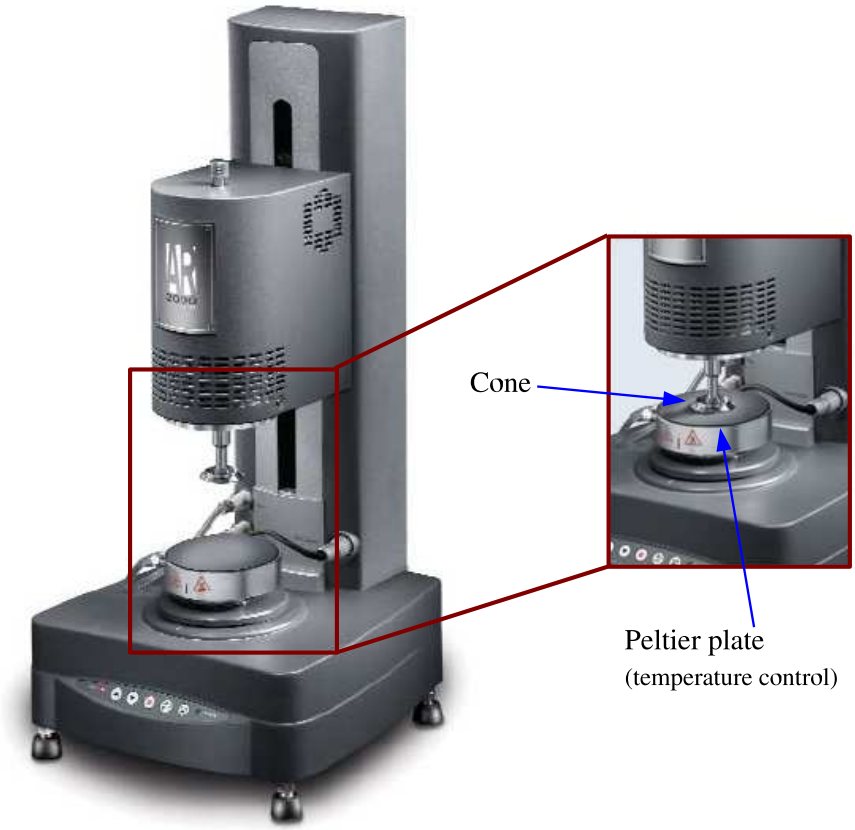
A major drawback of rheology for protein science is that large sample volumes are required, often greater than 500μl for a cone and up to millilitres for Couette cells. This puts a significant strain on stock protein solutions, which are typically purified in milligram quantities. Furthermore there is a limited number of experiments that may be performed on a sample before it begins to dry at the interface with air. This may lead to crystal particles forming in the solution and protein denaturation. Sample evaporation can be slowed by saturating the surrounding air with a suitable solvent but reliable experiments can only be performed on a sample for a few hours.

4.2 Circular Dichroism

Circular dichroism (CD) is a spectroscopic technique using circularly polarized light. If the electric component of an electromagnetic wave is confined to oscillate in one plane only, it is linearly polarized. Circularly polarized light occurs if the electric vec-



a)



b)

Figure 4.6: **a)** Schematic of a rheometer. The material (black) is between a plate (usually kept at a constant temperature) and cone. As the cone rotates it imposes a shear strain on the material. Note that the angle the cone makes with the plate,θ, has been exaggerated for clarity and is usually between 0°-5°. **b)** Photo of the AR2000 rheometer showing the cone and plate geometry. The sample is placed in the centre of the temperature controlled peltier plate and the cone lowered gently on top.

tor has a constant magnitude but rotates in the plane perpendicular to the direction of propagation. This is illustrated in figure 4.7 for right-handed circularly polarized light, which forms a right-handed helix as it propagates. A linearly polarized wave can be composed of two circularly polarized waves with identical magnitudes but opposite directions of rotation, as shown in figure 4.8. If left and right circularly polarized light are passed through a material and absorbed equally, the emerging beam is plane polarized. If, however, the material is optically active, one of the polarisations will be preferentially absorbed over the other (the effect known as dichroism) and the emerging beam is elliptically polarized, as shown in figure 4.8. Chiral molecules, which have no reflection plane and cannot be superimposed on their mirror image, also absorb right and left circularly polarized light to different extents, depending on their inherent chirality.

CD is defined as the difference in absorption between left and right circularly polarized light by an optically active, chiral sample [109]:

$$CD = A_l - A_r = \Delta A \tag{4.11}$$

CD spectra show the degree of ellipticity (Θ) and are reported by the spectrometer in millidegrees (mdeg). To compare spectra the *molar* ellipticity should be used (equation 4.12) as it takes into account the sample concentration (c, in g/ml), the cell path length (l, in mm) and the mean residue weight (w) [110]. The mean residue weight is 117 for ocr and around 110-115 for the majority of proteins.

$$molar\,ellipticity = \frac{w\Theta}{lc} \tag{4.12}$$

Chirality is an intrinsic feature of many biological and physical systems. The vast majority of biological molecules are chiral so will interact differently with right and left circularly polarized light. This makes CD a sensitive probe for investigating their three dimensional features. For example, CD can easily distinguish between A-, B- and Z-form conformations of DNA molecules. It is used to study binding or other subtle interactions between DNA, protein and ligand molecules. The principal application of CD, though, is predicting the secondary structure composition of proteins. This is particularly useful for proteins whose structure has not yet been characterised and is often done before the lengthy process of protein crystallography.

CD is commonly measured over two wavelength ranges: near UV (320nm-250nm) and far UV (260nm-190nm). At short wavelengths, absorption is predominantly due

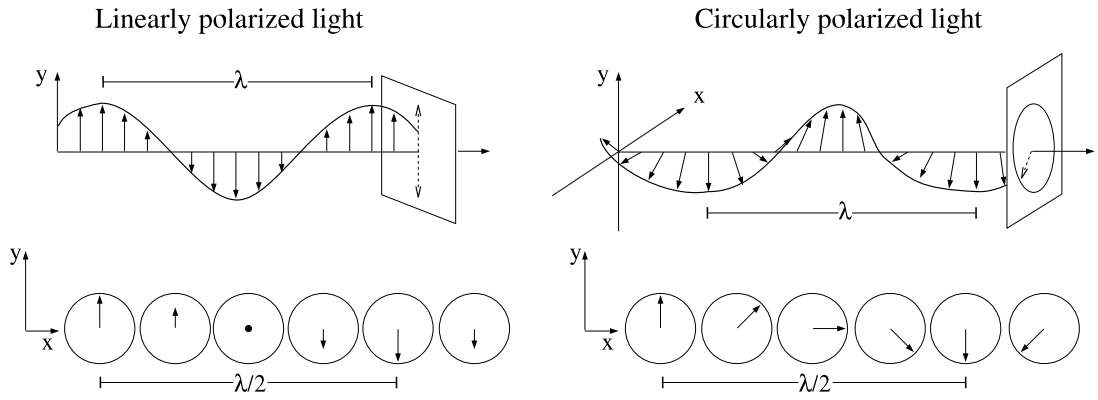


Figure 4.7: Diagram showing the electric component of linearly polarized light (left) and right-handed circularly polarized light (right). Under each diagram are the electric component vectors as viewed along the axis of propagation, showing the rotation of circularly polarized light.

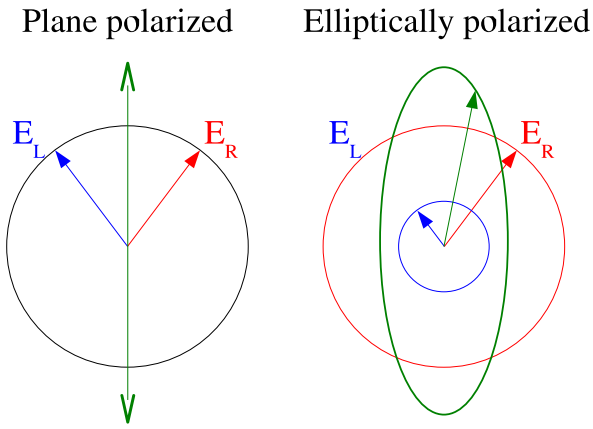


Figure 4.8: The combination of two circularly polarized waves rotating in opposite directions may produce a plane polarized wave if they have the same magnitude (left diagram) or an elliptically polarized wave if they differ in magnitude (right diagram). The latter may result if the circularly polarized waves are passed through an optically active material that preferentially absorbs left circular light.

to the peptide bonds in proteins. The CD spectra strongly depends on the proportions of α helices, β sheets or random coils present as they each absorb by different amounts at different wavelengths. By following changes in the far UV CD signal, we can study denaturing and refolding events. Absorption in the near UV is due to the aromatic side chains, in particular phenylalanine, tryptophan and tyrosine. CD signals at these wavelengths are sensitive to the local environment around the residues and can be used to track external events like ligand binding, in addition to folding.

Where experiments involving ocr are concerned, CD has a number of practical advantages over rheology. It requires only small sample volumes (150μl for far UV or 300μl for near UV) and low concentrations of protein (a strong gel is not required). CD spectroscopy is non-intrusive. We can probe the structure of the gel without the risk of destroying it and possibly even reuse the sample in another experiment. CD is also quick to measure; good quality results can be obtained within 30 minutes per sample.

A continuous flow of nitrogen gas is needed to remove oxygen from the CD spectrometer. There are two reasons why removing oxygen is vital. Firstly, ozone is produced when ultraviolet light hits oxygen. Ozone degrades the sensitive optical components. Secondly, it is most important to avoid oxygen at far ultraviolet wavelengths where it is highly absorbent. This reduces the available light for measurements, resulting in noisy signals[110]. Signal quality can be optimised by measuring a baseline to remove the buffer's contribution to the absorption spectrum and filtering buffers through 0.2μm filters to remove particles that could absorb or scatter the incident light. The absorption at far UV wavelengths is very sensitive to protein concentration so a shorter path length and more dilute solution is used compared to near UV, which needs a more concentrated sample [111].

4.2.1 Far UV

Each secondary structure element has a characteristic far UV CD spectrum, as shown schematically in figure 4.9. The structure of an unknown protein can be predicted by assuming its spectrum, $\Theta(\lambda)$, can be made up by a linear combination of the spectra of each secondary structure element, E_i, including some noise, N:

$$\Theta(\lambda) = \Sigma f_i E_i(\lambda) + N$$

where f_i is the fraction of the ith element. So the protein spectrum can be analysed

to give a percentage content of each secondary structure element. This principle has been expanded to use the spectra of a number of proteins whose structures are known to fit the measured spectrum. This allows a more accurate analysis as it includes contributions from other factors that affect the CD spectra, like β-turns and aromatic side chains [112].

Various methods exist for analysing far UV spectra. They differ primarily in the range of wavelengths used and in the number and type of reference proteins [113]. The most common analysis programs can be access through DichroWeb[2] [114, 115]. This service is provided by the BBSRC[3] Centre for Protein and Membrane Structure and Dynamics.

I used the K2D analysis program on the website as this only requires wavelengths down to 200nm. It uses a range of 18 proteins that have known crystal structures and CD spectra to determine the α-helix and β-sheet composition of the sample. The predicted secondary structure is built up from the reference proteins by combining them in different ratios. The relative weights are refined in an iterative process to achieve the best fit for the predicted CD spectrum to that of the protein spectrum [116, 117].

Buffers that absorb strongly in the UV region should be avoided, for example sodium chloride where the chloride ion absorbs very strongly below 200nm. Tris-HCl is among the most useful buffers, including phosphate, as it has a low absorbance in CD spectra [53]. The lower the wavelength cut off, the more accurately the spectra can be analysed. For example, by using wavelengths down to 175nm results can be achieved that are 95% in agreement with secondary structure values as determined from x-ray crystallography. Practically, however, it is difficult to collect high quality data below 200nm [110]. Here the system becomes increasingly sensitive to oxygen, and both samples and buffers absorb more at lower wavelengths so spectral noise increases.

4.2.2 Near UV

The tertiary structure of the protein can be probed by measuring in the near UV region. Here the aromatic amino acid side chains are key. They are intrinsically achiral so the

[2] (*http://www.cryst.bbk.ac.uk/cdweb/html/*)
[3] Biotechnology and Biological Sciences Research Council

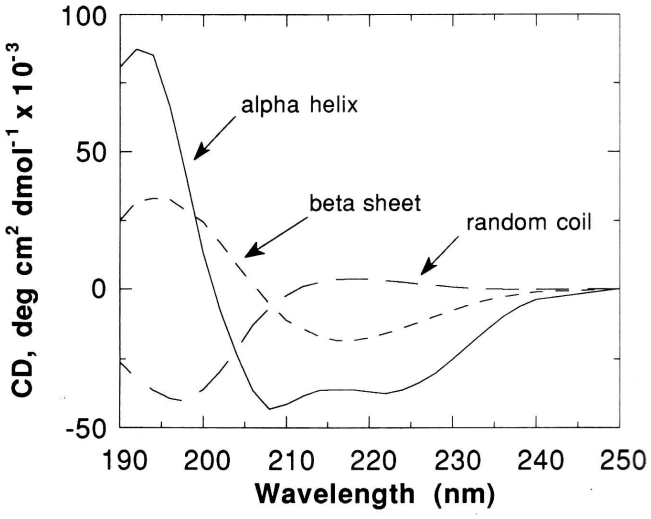


Figure 4.9: Schematic of characteristic far UV circular dichroism spectra for the secondary structure elements: alpha helix, beta sheet and random coil. Measured spectra are analysed in terms of these building blocks. Diagram reproduced from [53] and based on the spectra of poly(lysine) in the respective conformations.

intensity of the signal is very dependent on their environment and orientation.

CD theory is not advanced enough to allow detailed analysis of near UV spectra, which is influenced by hydrogen bonding, polar groups and polarizability among other factors. In particular the mobility of the side chains has a significant effect, where the less mobile the side chains the higher its intensity [110]. Under 3.2M ammonium sulphate ocr has a fairly typical near UV spectrum. Above 3.2M the extra salt confers chirality to the system.

4.2.3 Method

A JASCO J-810 Spectropolarimeter connected to a Peak Scientific nitrogen generator (NG9-3LA) was used to measure the CD spectrum. The experiments were carried out at 20°C with a band width of 1nm, a 2s response, a 0.2nm data pitch and with a sensitivity of 100mdeg. They were performed at 20nm/min over 5 accumulations. A Peltier plate was used in the sample chamber to ensure a constant temperature throughout the experiment. An ocr concentration of 30μM ocr was used in the gels. The sample was measured in a 10mm cuvette for near UV (320nm-250nm) and a 1mm cuvette for far UV (260nm-190nm). The cuvettes were made specifically for CD measurements

by Starna and sealed during the scans to prevent evaporation. Each experiment was repeated with the respective salt buffer alone under identical conditions to obtain a baseline, which was subtracted from the gel's spectrum to produce the final CD result. The baseline and sample were measured using the same cuvette as it may have small birefringent effects that need to be taken into account.

4.3 Atomic force microscopy

Atomic force microscopy (AFM) is a method of probing surfaces on an atomic scale. It is capable of producing images with a horizontal resolution of better than 10nm and a vertical resolution in the angstrom region. It can also measure forces down to 10pN [118].

Originally used for imaging materials like metals, semiconductors and catalysts, the field rapidly expanded. It is now regularly used for sensitive measurements of elasticity, chemical bonding and electrostatic, magnetic and capillary forces. Biological systems are readily analysed by AFM, where the forces needed to separate macromolecules, stretch polymers and rupture fibrils can be measured. For example AFM has been used to unfold Titin, a large protein that controls the elasticity of striated muscle [119] and probe the formation of amyloid fibrils, some of which are associated with diseases like Alzheimer's and CJD [120]. High-resolution three-dimensional imaging has also enabled *Eco*RI restriction sites on bacteriophage λ DNA to be mapped and DNA condensation into toroids to be observed [119, 121]. The process of restriction, as described in section 2.2.1, has even been observed in real-time (figure 4.10). An *Eco*KI restriction-modification enzyme bound to the plasmid pRH3 was seen to reel the DNA in, by translocation, over a period of a few minutes. The enzyme remained stationary throughout the process and after 7 minutes fragments of DNA were observed diffusing away from the *Eco*KI complex. A further 7 minutes later almost all the DNA had been cleaved, leaving only few small fragments bound to the enzyme. This process is slightly different from the one described previously, which required two R-M enzymes. Here, the single enzyme dimerises soon after binding to the DNA, allowing translocation and cleavage between two bound *Eco*KI molecules.

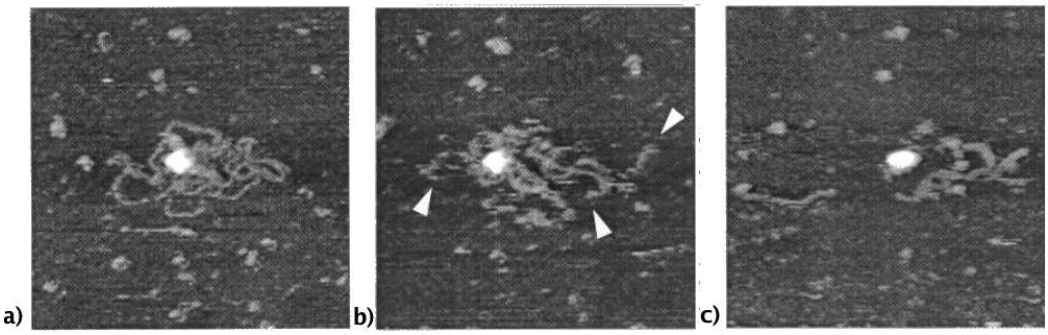


Figure 4.10: AFM images showing the cleavage of the DNA plasmid pRH3 by the restriction-modification enzyme *Eco*KI. **a)** The large white spot is the bound *Eco*KI, surrounded by loops of DNA. **b)** These are gradually cleaved and diffuse away from the *Eco*KI complex (as indicated by the white arrows). **c)** Eventually only a small core of DNA fragments remain bound to the enzyme. Reproduced from [121].

4.3.1 Method

A detailed topographic map of surfaces is produced by scanning a fine tip gently across the sample. The tip is attached to the end of a cantilever, which allows the tip to follow the elevation of the surface, like a stylus on a record player. A laser reflecting off the cantilever is used to detect its deflection, as shown in figure 4.11. Signals from the photodiode are passed back in a closed feedback loop to a sensitive piezoelectric scanner that continuously adjusts the height of the sample. In this way the tip does not collide with the surface and the cantilever deflection, along with the force imposed on the sample, is kept small. The piezo also controls the scanning in the x and y directions [122, 118, 123].

The surface structure can be measured in two ways. In contact mode the tip traces along the top of the sample. A constant force between the tip and the surface is maintained by the piezo changing the sample elevation. This is converted into surface height information and a relief image produced where the elevation is represented by a monochrome gradient (*e.g.* highest areas lightest, lowest areas darkest). However, the lateral force may cause sample distortions and result in images that appear to have been swept across by a brush. This can be avoided by using tapping mode. Here the cantilever is oscillated near its resonant frequency (*e.g.* 200kHz in air) and gently taps the sample. As it encounters different surface heights the oscillation amplitude changes. The piezo is used to compensate for these changes and maintain a constant

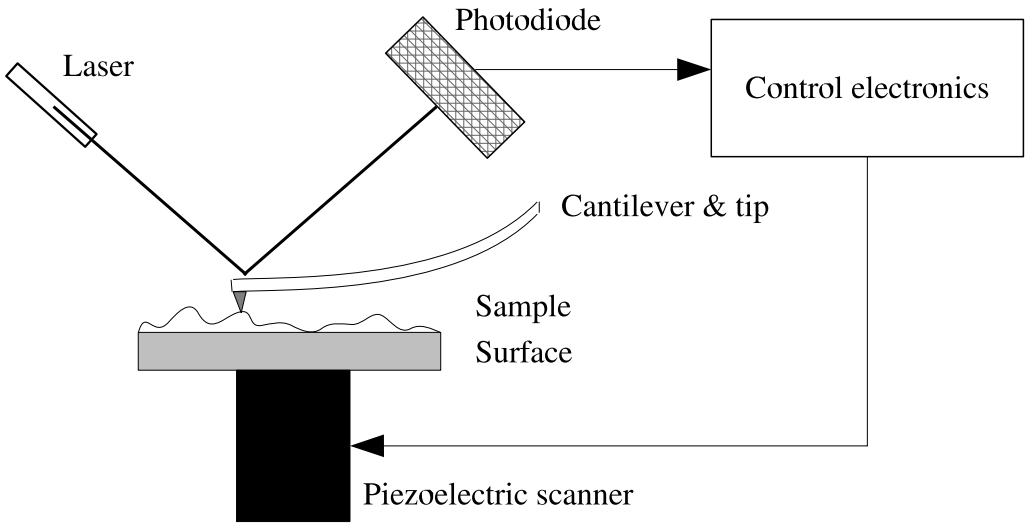


Figure 4.11: Diagram of an atomic force microscope. The vertical displacement of the tip, as it moves across the surface of the sample, is detected through variations in the laser signal reflecting of the top of the cantilever. A closed feedback loop from the photodiode to the piezoelectric scanner adjusts the height of the sample to ensure the cantilever deflection is kept within its optimal range.

amplitude of cantilever oscillation. Once again the change in sample elevation provides a measure of the surface topography.

The sample of interest is placed on a clean, flat, hard surface. The most common substrates are glass or mica. Mica is an aluminium silicate crystal that is atomically flat and can achieve a vertical resolution of less than a nanometer. Glass is cheaper but has a surface roughness of 1-2nm so is best used for larger structures. The success of samples adhering to the surface depends on their hydrophobicity and charge. Glass and mica are both negatively charged but can be coated to give alternative properties. For example, a poly-L-lysine coating gives a positively charged surface.

The quality of the image, in particular the resolution in the x-y plane, is limited by the size of the tip. If the end of tip is significantly larger than the structures the image may reflect the shape of the tip more than the molecules. The tip may also pick up part of the sample so you are no longer imaging with a defined point. This is revealed by surface features having the same orientation over the whole image. If an unusually large structure is imaged the Piezoelectric scanner may not react fast enough when

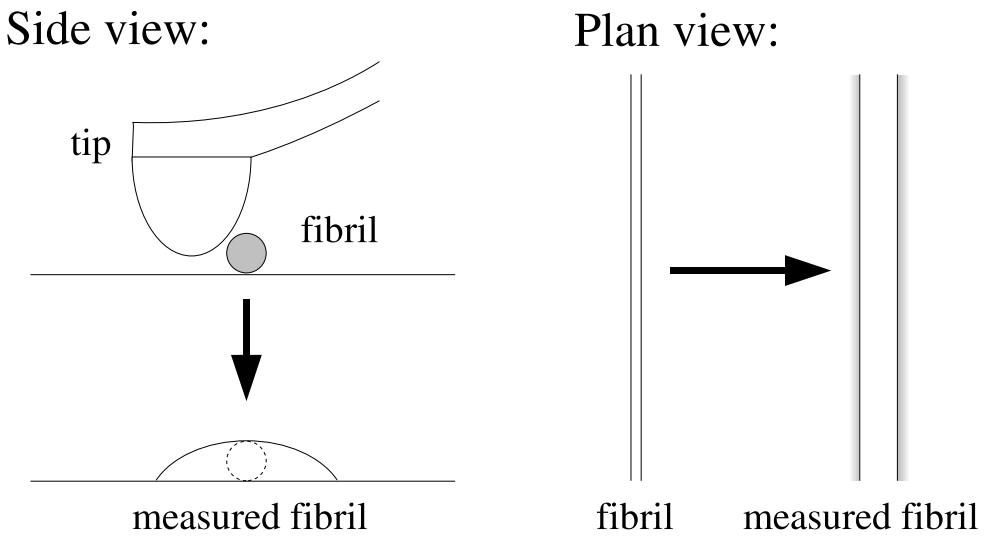


Figure 4.12: An illustration of the effect of convolution on a fibril being measured with an AFM tip. The fibril width appears wider in the x-y plane than in reality due to the much larger size of the tip. Vertical height measurements are more accurate as they rely only on the vertical displacement of the tip, as measured by the laser.

returning to the rest of the sample and shadows may be produced beside the structure in the image [122]. Figure 4.12 illustrates the effect known as tip convolution. The tip is not an infinitely sharp point, but a rounded bowl, which defines the smallest scale measurable in the x-y plane. The effect of the convolution was estimated by Andreas Engel *et al.* [124]: a particle of diameter D, measured with an AFM tip of radius R, exhibits a width at half height of

$$W = 2\sqrt{RD + \frac{D^2}{4}} \tag{4.13}$$

Resolution in the z-axis, and therefore measurements of heights, is most accurate due to the laser recording the vertical movement of cantilever and tip.

4.3.2 Ocr and AFM

A Multimode AFM with a Nanoscope IIIa controller and "E" scanner from Veeco was used to image ocr gels. Triangular tips with a full tip cone angle of 30°, a radius of less than 10nm and height of 15-20μm were attached to silicon cantilevers from μmasch. The resonant frequency of the tip is between 120-190 kHz and its spring constant is

between 2.5-8.5 N/m. The "E" scanner has a range of 0-15μm. To reduce vibration the equipment was mounted on a heavy, marble table. Data was recorded in both trace and retrace directions simultaneously, along with the raw amplitude to check for scan artifacts. Scan rates of 1 Hz, 1.38 Hz and 1.91 Hz were used for 5μm, 3μm and 1μm scan sizes respectively. Tapping mode was employed with the maximum resolution of 512 points per line for recording images. Samples were made with a range of salt and ocr concentrations to probe the gel transition. 10μl of each sample were deposited on glass surfaces, dried overnight at room temperature (under a cover to prevent dust settling on top) and imaged within 24 hours of deposition. No AFM studies were performed in the hydrated state. This would have avoided concerns about drying the samples, like salt crystals forming or fibrils collapsing, but there was only time for one form of AFM experiments. "Wet" AFM was avoided because it required larger sample volumes, it wasn't clear if the fibrils in the gel would adhear to the glass surface and there were concerns about the shear forces imposed on the gel when filling the sample holder.

4.3.2.1 Calibration

AFM cantilevers cannot be manufactured to have a precise resonant frequency. Rather, for each type, it lies within a certain range. This is quoted along with the corresponding spring constants. The resonant frequency in air of a new cantilever must be measured and the equivalent spring constant calculated. For example figure 4.13 shows a cantilever measured at 158 kHz compared to the quoted values to give a spring constant of 5.3 N/m. Once the cantilever has been calibrated the oscillation frequency of the tip is set just below the peak value to avoid artifacts from oscillating at the resonant frequency.

Distances in the x-y plane and in the z-axis were calibrated using a standard grid of square wells with a depth of 200nm and periodicity 10μm (see figure 4.14).

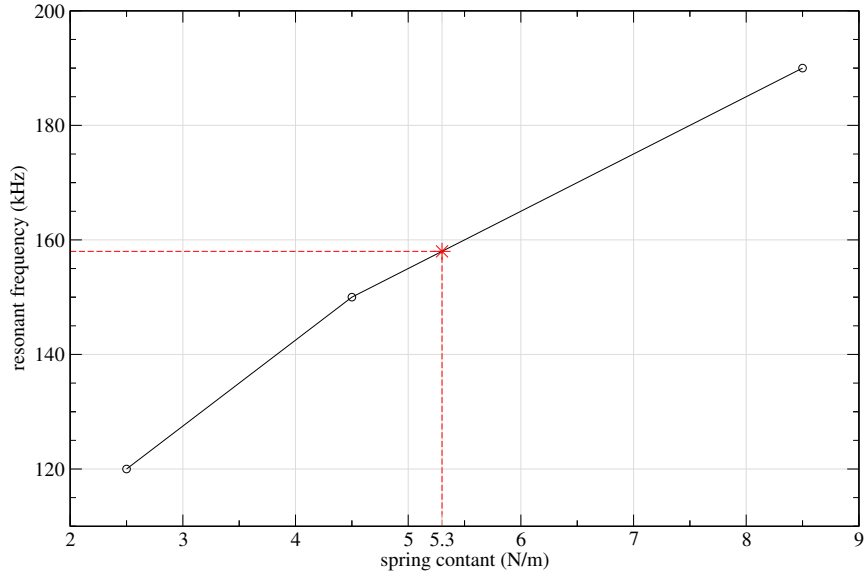


Figure 4.13: The black line shows the AFM cantilever's quoted frequencies and corresponding spring constants. The measured resonant frequency of the chosen cantilever was 158 kHz, corresponding to 5.3 N/m (red, dotted line). This calibration check was performed for each new cantilever.

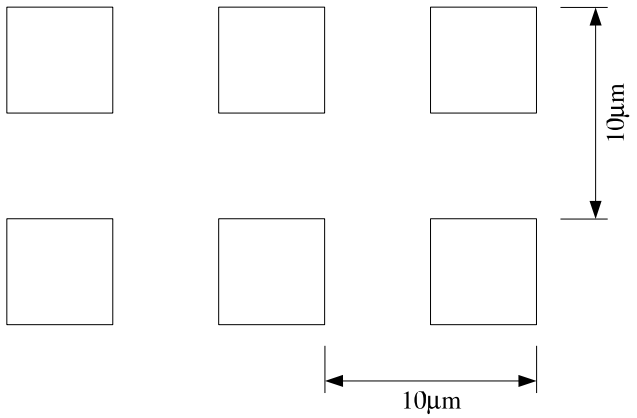


Figure 4.14: Schematic of the calibration grid. Each square well has a depth of 200nm and a separation of 10μm.

Chapter 5

Results

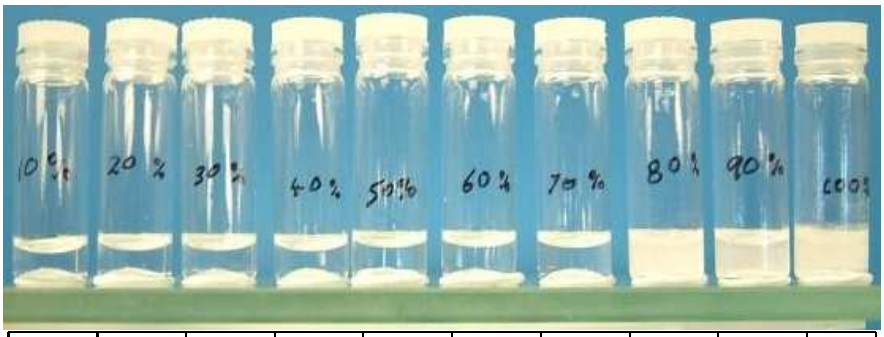
5.1 The ocr gel

5.1.1 Preliminary observations

To form a gel, solid ammonium sulphate may be added to an ocr solution. The transition is extremely rapid and occurs suddenly at 80% saturated ammonium sulphate (3.2M)[1]. To illustrate this, ten solutions of 1mg/ml (36μM) ocr in 20mM Tris-HCL, pH 8 were made up at room temperature and solid ammonium sulphate added to the required concentration (see table 5.1). The resulting solutions are shown in figure 5.1. Below 80% ammonium sulphate saturation the salt dissolves completely and the protein does not precipitate, leaving a transparent, clear solution which flows like water. There is a sudden change at 80% saturation and above. The salt dissolves completely but the samples appear gel-like and trap the air released by the salt dissolving.

A number of other samples were prepared, varying in buffer strength (0 to 1M Tris-HCl, pH 8, with 1mg/ml ocr) and ocr concentration (0.1 to 5 mg/ml ocr in 20mM Tris-HCl, pH 8). Each formed a gel in 80% ammonium sulphate. A crude observation was made of their viscosity by tilting the samples 90° (fixed to a plate) to see how quickly they flowed. Figure 5.2 shows a photo from each experiment, taken as the gels flow down the sample bottles. The head position of the samples has been highlighted with a blue line. As expected, the flow rates for each buffer strength appear very similar

[1] Initial experiments recorded salt concentrations in percentage saturation, and protein concentrations in mg/ml, as these units are commonly used by molecular biologists. 80% saturated ammonium sulphate is equivalent to 3.2M ammonium sulphate, and 1mg/ml ocr is equivalent to 36μM ocr. Molar quantities will be adopted once the initial observations have been described.



0.4M	0.8M	1.2M	1.6M	2.0M	2.4M	2.8M	3.2M	3.6M	4M
←—Liquid							Gel —→		

Figure 5.1: A series of 36μM ocr solutions in 20mM Tris-HCL, pH 8, with solid ammonium sulphate added to increasing concentrations. The molar concentration of ammonium sulphate in each solution is shown in the table and the equivalent saturation is written on the bottles. 100% saturation equates to 4M ammonium sulphate. The rapid gel transition occurs at 80% saturation (3.2M). The gels appear opaque due to air trapped in the viscoelastic fluid.

Molarity	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2	3.6	4.0
Saturation	10	20	30	40	50	60	70	80	90	100
g added to 1L	55	113	176	242	314	390	472	516	657	761

Table 5.1: Equivalent concentrations of ammonium sulphate: molar, saturation and grams to be added to 1L of buffer solution.

(a), and clearly the concentration of Tris-HCl used has no effect on the viscosity of the gel. Increasing the concentration of ocr (b), however, markedly increases the viscosity of the gel. This, too, was anticipated since more protein should result in more bonds to form a tighter network in the gel. For the samples with increasing salt concentration (c), no difference in flow rate was observed until 80% saturation, where a gel forms. Here, the viscosity suddently, and significantly, increases. Surprisingly, above 80% saturation, the viscosity begins to decrease. The gel appears to be weakening. This visual observation was to be corroborated later by other experiments.

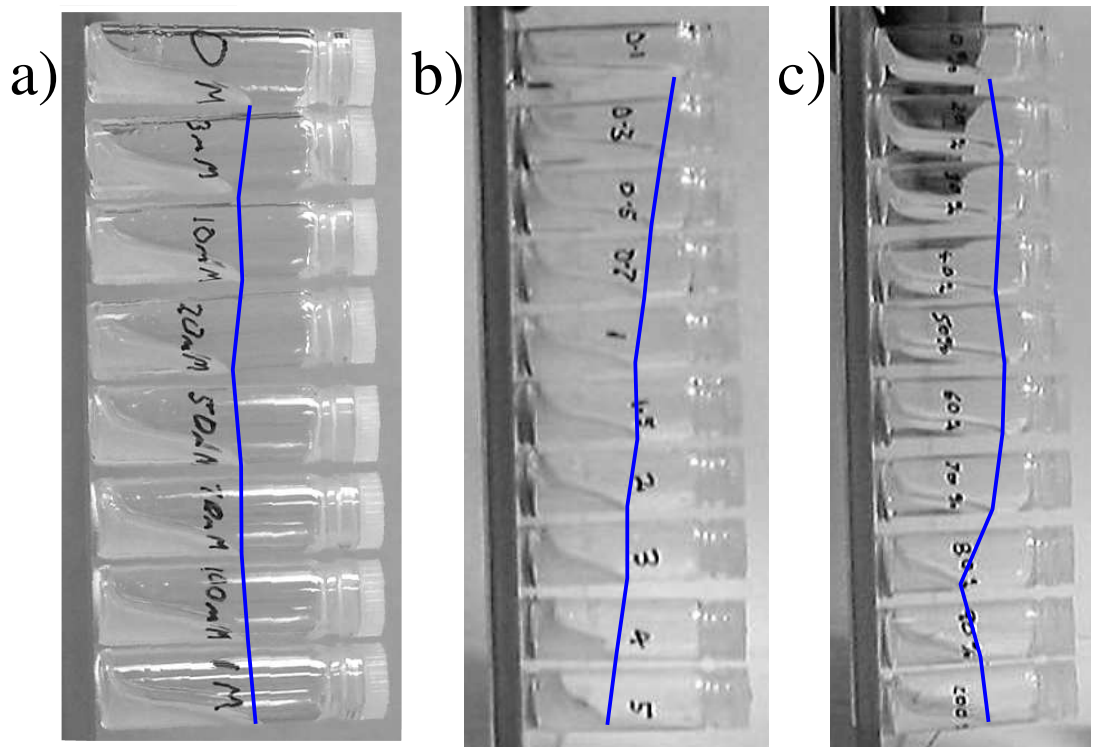


Figure 5.2: Visual observations of the viscosity of ocr gels under **a)** varying buffer strength (0 - 1M Tris-HCl, pH 8) for 1mg/ml ocr in 80% ammonium sulphate, **b)** varying ocr concentration (0.1 - 5mg/ml) in 20mM Tris-HCl, pH 8, 80% ammonium sulphate, and **c)** varying ammonium sulphate concentration (0 - 100% saturation) for 1mg/ml ocr in 20mM Tris-HCl, pH 8. Samples were prepared in bottles fixed to a horizontal plate, which was then tilted vertically to observed the viscosities. The head position of each sample is highlighted by the blue line. Buffer strength has no effect on the viscosity of the gel; increasing ocr concentration results in increasing viscosity; a sudden increase in viscosity is observed at 80% ammonium sulphate but this decreases above 90%.

Figure 5.3: A *transparent* gel formed by 36μM ocr after dialysis against 3.3M ammonium sulphate (82.5% saturation). Both were suspended in a 20mM Tris-HCl buffer at pH 8. The dialysis was performed at 4°C. Note that the bubbles were deliberately created to illustrate the gel's retention properties.

5.1.2 Gel formation process

Once gelled, the air trapped in the solid ammonium sulphate, which is released upon dissolving the salt, was consequently trapped in the gel. Optical and rheological experiments are hampered by these bubbles. Attempts were made to remove air from the salt prior to its use, by grinding it down to a fine power, drying it in an oven and storing it in a desiccator, but to no avail. An alternative method was tried where aqueous salt solution is added to a concentrated protein sample but this did not form a homogeneous gel due to imperfect mixing. The most effective procedure discovered is to dialyse a protein sample against a salt solution of the required concentration. This results in a uniform, transparent gel, free from bubbles and contaminants. An example gel formed by dialysis is shown in figure 5.3. If this gel is dialysed in a 20m Tris-HCl buffer (pH 8) without salt it returns to a simple protein solution. It can then be dialysed once again in a salt solution to remake the gel.

Ammonium sulphate has a molecular weight of 132.14 and at room temperature a

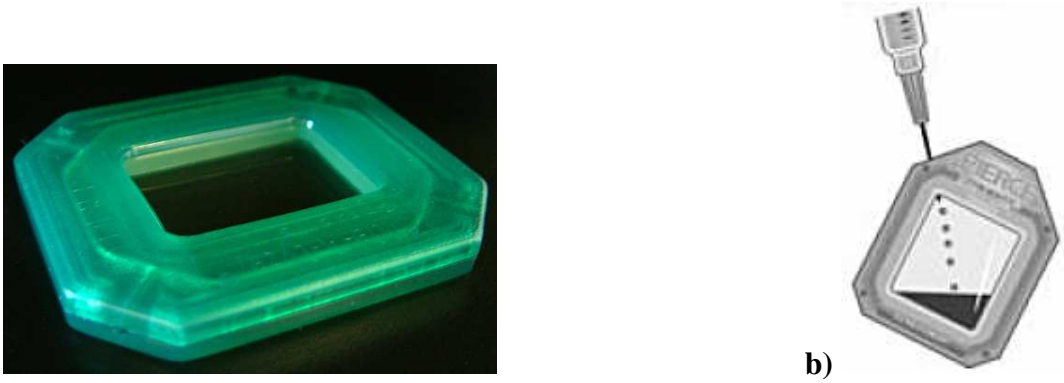
saturated solution has a concentration of 4M (1.77g/cc) [94]. The majority of experiments on ocr gels were performed at a salt concentration of 3.3M. When making up salt solutions care has to be taken with the final volume. For example to make up a 3.3M solution 175.35g is added to 300ml buffer, giving a final volume of 400ml and final concentration of 438.4g/L = 3.3M (see table 5.1).

The following procedure describes how a 100μl sample of 1mg/ml (36μM) ocr, 3.3M ammonium sulphate, 20mM Tris-HCl, pH 8 solution was made. 2.5μl of stock ocr solution (40mg/ml in 20mM Tris-HCl, pH 8) was added to 97.5μl of 2M ammonium sulphate, 20mM Tris-HCl, pH 8 solution. 2M ammonium sulphate was used rather than a pure buffer solution, to minimise any volume changes in the sample during dialysis against high salt concentrations, while being far enough away from the critical gelling concentration not to affect the process.

The standard dialysis method of placing the sample in dialysis tubing and sealing the ends is only practical for large volumes as recovering small volumes from the folds of the tubing is neither efficient nor 100% successful. Dialysis cassettes, made by Pierce, reduce sample loss and may be used for volumes down to 0.5ml (figure 5.4). However, they are expensive and are only single-use. Instead, a reusable "cassette" was made from the top of a glass bottle. The bottom of a 15mm diameter glass bottle was sliced off, leaving the plastic lid and the top 1cm of the glass (as shown on the right of figure 5.5a). A micro dialysis cassette may be fashioned in a similar manner from an eppendorf tube, by cutting off the bottom 5mm below the rim (figure 5.5a). The sample is placed in the lid of either, covered with a single sheet of dialysis tubing and secured in place by the remainder of the bottle or eppendorf (figure 5.5b). Sample volumes of 150μl and 1.5ml may be used for the eppendorf and bottle tops respectively. 10000 MWCO SnakeSkin dialysis tubing from Pierce was used for preparing ocr gels. The sample was dialysed against a 3.3M ammonium sulphate, 20mM Tris-HCl, pH 8, solution. Dialysis was performed at 4°C for 3 hours. The final protein concentration can be checked by UV spectroscopy.

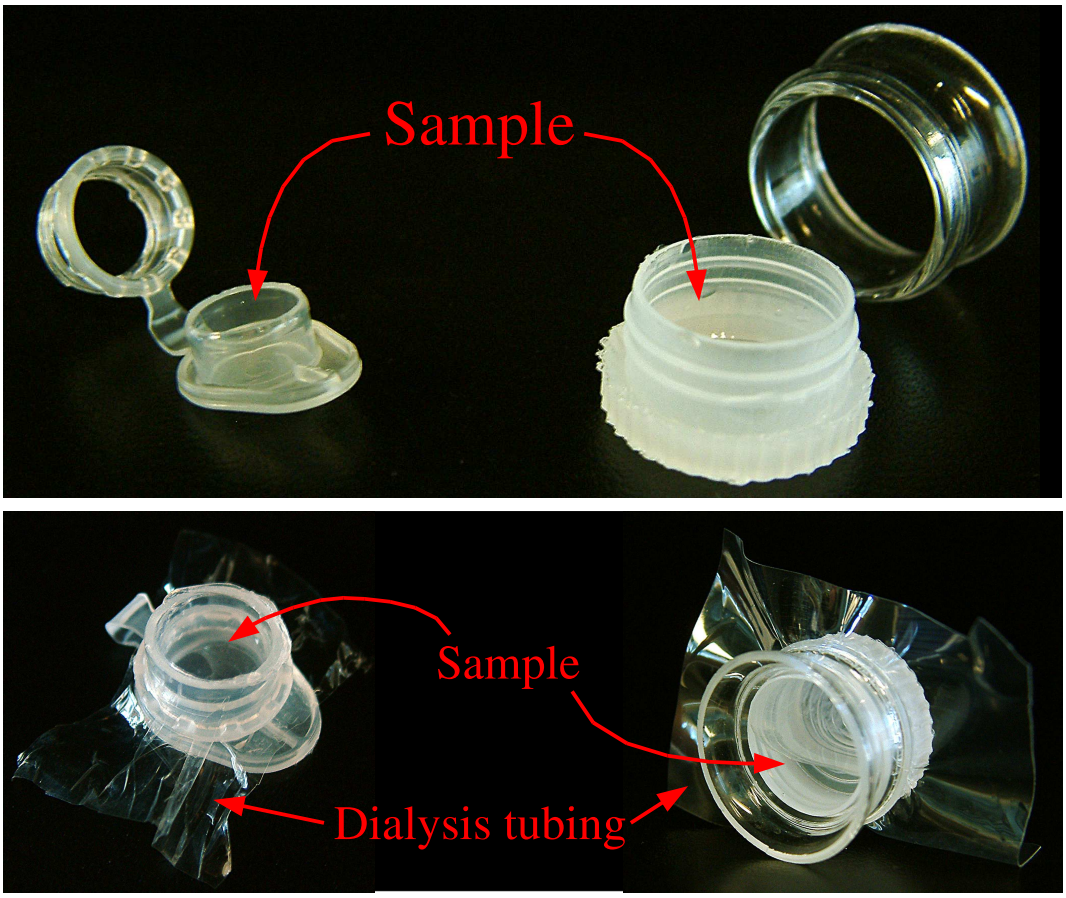
5.2 Rheology

In the experiments concerning ocr the temperature was controlled by a peltier plate and fixed at 20°C. A 40mm (diameter), 2°cone and plate geometry was used, requiring a



a) **b)**

Figure 5.4: **a)** Dialysis cassette made by Pierce. **b)** The sample is injected into a cavity bounded by two dialysis tubing windows (image reproduced from instruction leaflet).



a)

b)

Figure 5.5: Two “Blue Peter” dialysis cassettes. The one on the left is an eppendorf tube and the one on the right is a small glass bottle, both with their bottoms cut off. The sample is deposited in the lid **(a)**, dialysis tubing placed over the top and fixed in place by the remainder of the eppendorf or bottle **(b)**.

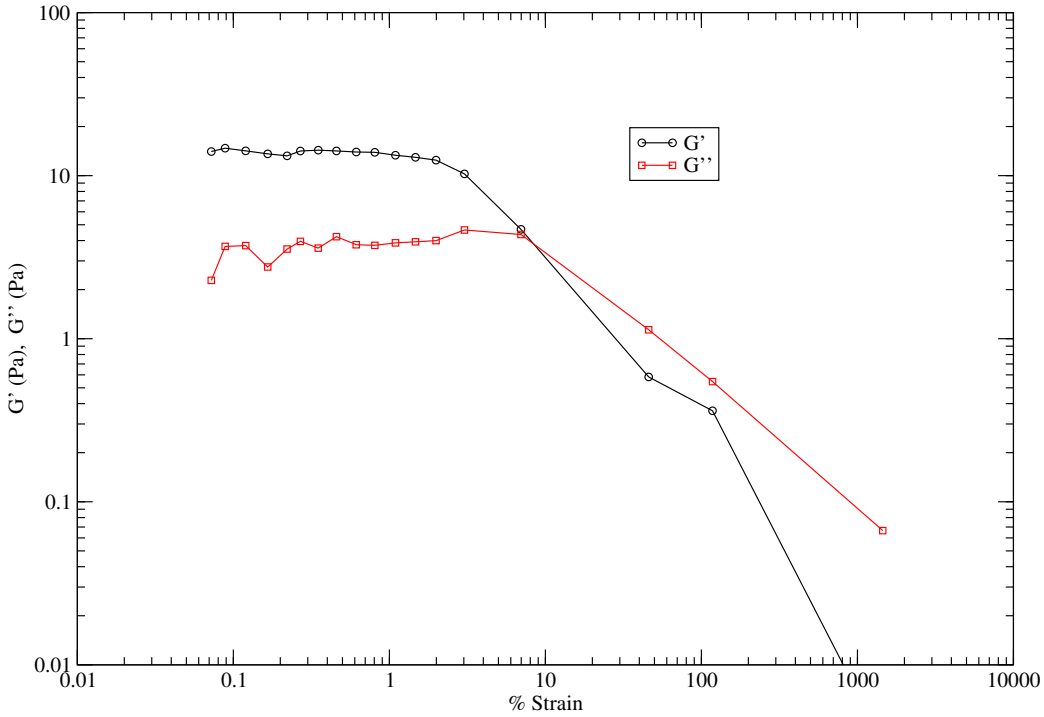


Figure 5.6: The response of an 18μM ocr in 3.3M ammonium sulphate buffer solution to a strain sweep at a fixed frequency of 0.5rad/s. The linear viscoelastic region is clearly defined by the critical strain limit at $\sim$2% (equivalent to 0.1Pa).

sample volume of 600μl. The protein gel was allowed to equilibrate at room temperature before placing on the plate and a saturated salt solution was used in the cone's solvent trap to reduce evaporation.

5.2.1 Oscillatory stress

For oscillatory probes, experiments must be performed in the linear viscoelastic region so that the structure remains intact. Ocr gels were made with 3.3M ammonium sulphate in 20mM Tris-HCl, pH 8, and with four concentrations of ocr: 18μM, 36μM, 72μM and 144μM (corresponding to protein volume fractions of 0.04%, 0.08%, 0.16% and 0.32%). An oscillatory torque was applied to each sample over a range of stresses, at a frequency of 0.5rad/s. Figure 5.6 shows the linear response of the storage and loss moduli for the weakest gel, 18μM ocr, until the critical strain of $\sim$2% (equivalent to 0.1Pa) is approached. Above this, G' and G'' become dependent on the stress. The bonds between structure elements are broken and the moduli diminish.

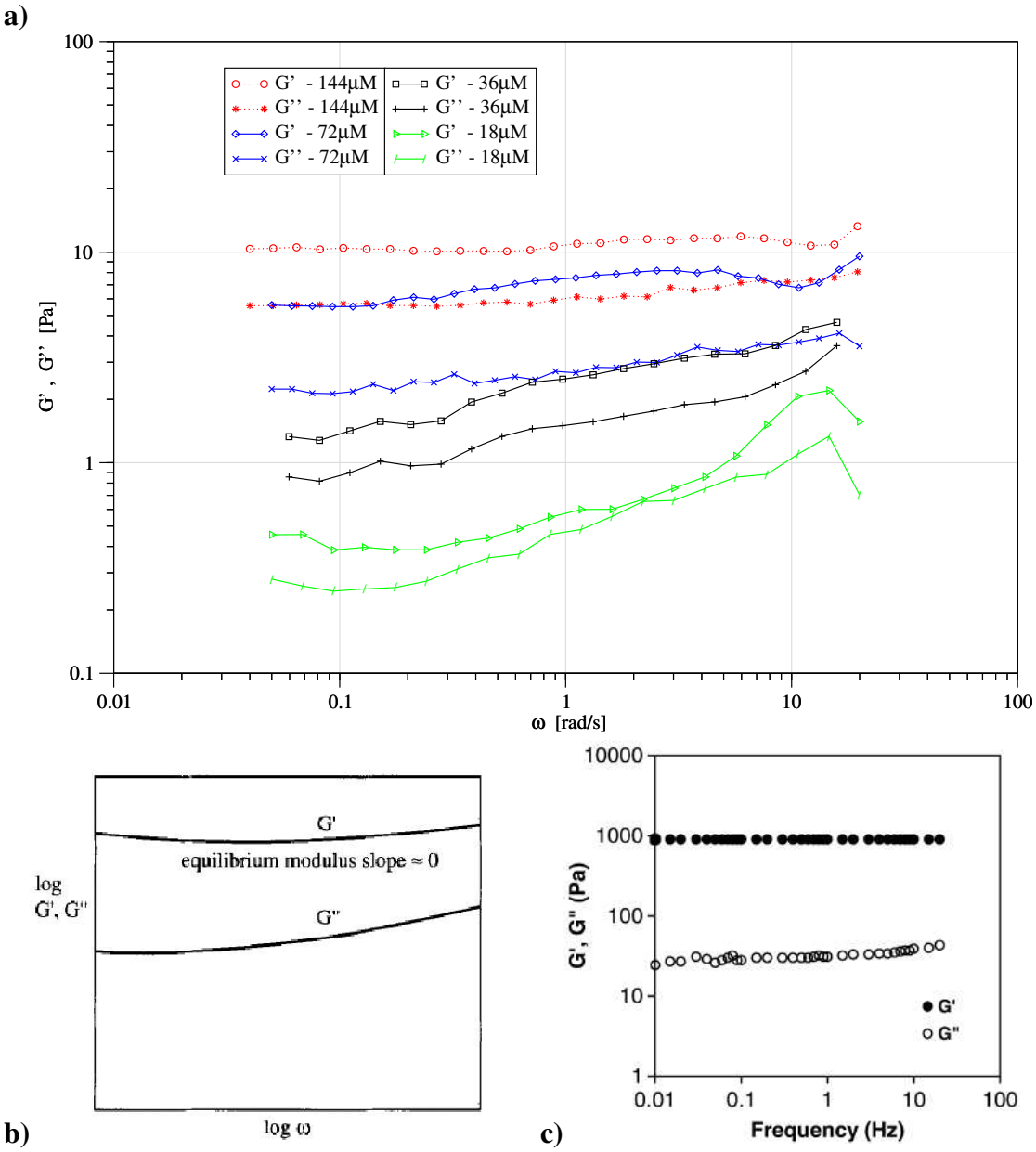


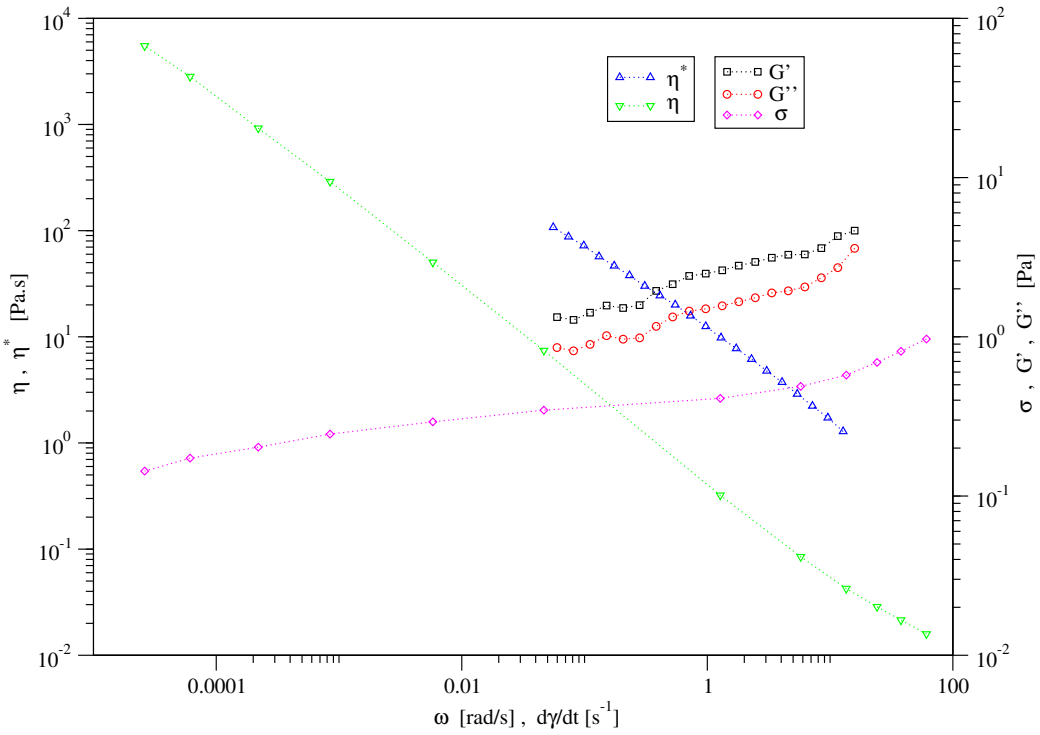
Figure 5.7: **a)** Rheology of 18μM (green), 36μM (black), 72μM (blue) and 144μM (red) solutions of ocr in a 20mM Tris-HCl (pH 8.0) buffer containing 3.2M ammonium sulphate. The storage ($\triangleright, \square, \Diamond, \bigcirc$) and loss ($/, +, \times, *$) moduli were measured using an oscillatory stress of 0.1Pa. **b)** A typical dynamic mechanical spectrum for covalently cross-linked networks, as predicted by G.Kavanagh and S.Ross-Murphy (figure reproduced from: [11]). **c)** Rheology of a 5% polyacrylamide gel at 25°C, giving a storage modulus 100 times greater than that of ocr gels (figure reproduced from [125]).

Performing a frequency sweep within the linear region is probably the most common form of rheological testing. The amplitude of the probe (the stress) was fixed at 0.1Pa while the frequency was decreased from 40 to 0.04 rad s^{-1}. Figure 5.7a) shows the storage and loss moduli for a range of ocr concentrations in 3.2M ammonium sulphate. Here G' is larger than G'' over the entire frequency range examined. Ocr clearly behaves as a gel in this region of oscillation [126], albeit a weak gel; a 144μM ocr gel only achieves a shear modulus of G=10Pa, a factor of 100 lower than a 5% polyacrylamide gel [125]. The parallel behaviour of G' and G'' has been seen in polymer gels at their percolation threshold [102]. Figure 5.8a) shows a summary of both the oscillatory and steady shear rheology results for a 36μM ocr gel. The dynamic viscosity η^*, obtained from oscillatory measurements, is always greater than the steady shear viscosity η. This indicates that the structure is weakened by large strains during steady flow.

Similar behaviour has been observed with a number of proteins. Figure 5.8 b) shows the rheology results from a β-lactoglobulin gel formed at high temperature[127]. Heating causes proteins like β-lactoglobulin, collagen and BSA to partially unfold [128]. The rod-shaped hear-denatured collagen molecules refold to form fibrils on cooling, while on heat-denatured lactoblobulin, groups are exposed that may form cross-links by bonding covalently, like cysteine groups that form disulphide bonds, or non-covalently, like buried hydrophobic regions. The proteins cluster into disordered clumps or flocs, which commonly have a self-similar or fractal anatomy. They aggregate into polymers, forming random cross-links with each other until they span the liquid volume, whereupon a continuous three-dimensional network or a gel is formed [129, 126, 130, 131]. The protein actin and the polysaccharide Xanthan, on the other hand, form gels via entanglement networks[132, 133]. Individual molecules polymerize into a double-stranded helical filament. Polymer strands cross and loop around each other to form a semiflexible network that behaves like an elastic solid. It is thought the protein lysozyme simply aggregates into small lumps that in turn join to make larger fractal structures

Ocr may form a gel by either of these processes or through an entirely different means. It has been suggested that storage and loss moduli following the trend in figure 5.7b) indicates a gel formed by chemical cross-links rather than entanglement networks [11]. However, rheology alone is not enough to distinguish between them.

a)



b)

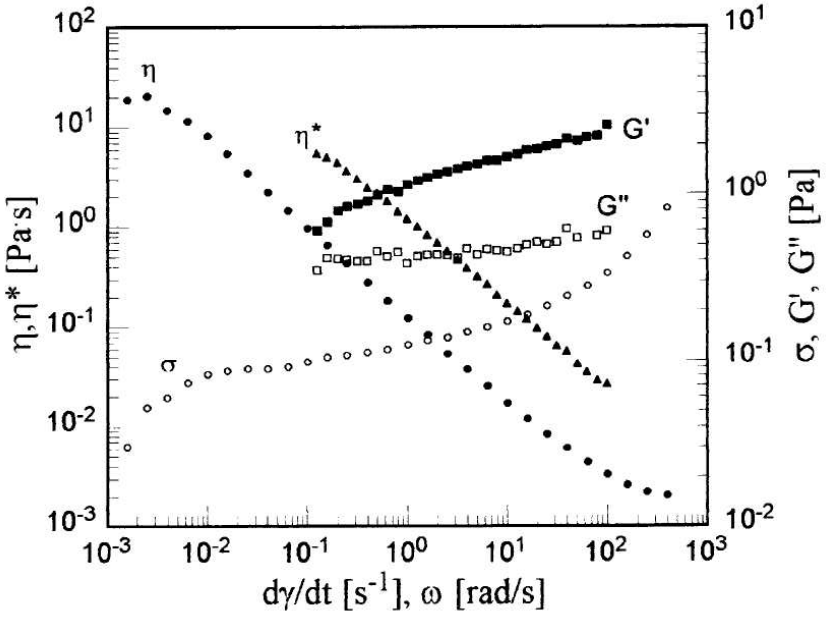


Figure 5.8: **a)** Rheology of a 36μM ocr solution in a 20mM Tris-HCl (pH 8.0) buffer containing 3.3M ammonium sulphate. The storage (□) and loss (◯) moduli and dynamic viscosity (△) were measured using an oscillatory stress of 0.1Pa. The steady shear stress (◊) and viscosity (▽) from steady shear flow experiments are also shown. **b)** Similar results obtained by *Ikeda and Nishinari* [127] for the protein β-lactoglobulin (5% w/w) in a gel formed by heating to 70°C.

Furthermore ocr is markedly different from the proteins considered above. It has a characteristic shape and specifically located charges. It also forms a gel at high salt concentrations; not high temperatures. In fact ocr is very stable at high temperatures, only fully denaturing at around 80°C [55]. Nor does it denature in the presence of salts such as ammonium sulphate.

5.2.2 Shear flow

A simple method for probing an unknown material is to apply a steady shear flow. Figure 5.9a) shows a steady shear experiment where the stress was decreased from 1Pa. Six solutions of 100μM ocr were prepared with different concentrations of ammonium sulphate in 20mM Tris-HCl, pH 8. The 0.0M sample behaves as a Newtonian liquid with a constant viscosity (fixed gradient). The solutions containing salt have three distinct regions: at low and high rates the flow is Newtonian but in the middle region the viscosity decreases and shear thinning is observed. This behaviour is commonly observed in polymer melts [103]. At lower shear rates the sample is able to flow whilst maintaining its overall structure but as the shear rate increases the structure can no longer bear the increased stress and breaks up.

There is also a trend across the different salt concentrations, which is highlighted in figure 5.9b). Here three stresses, each of a different order of magnitude, are chosen and the strain rate across all the salt concentrations plotted. There is a clear decrease in strain rate at $\sim$2.8M to a minimum at 3.3M, after which the system partially recovers. This corresponds to an increase in yield stress at the gel point of the sample and reflects results obtained by circular dichroism (figure 5.14b) and particle tracking (figure 6.4). Both show a sharp transition around 3.0M ammonium sulphate and an incomplete recovery.

5.2.3 Creep

A more sensitive measure of yield stress was acquired through creep experiments. A range of stresses from 0.01 to 1.0 Pa were applied to a 100μM ocr sample in a 3.3M ammonium sulphate, 20mM Tris-HCl buffer, pH 8 (figure 5.10 **a)**. During equilibrium a power law of the form $y = Ax^B$ was fitted to the final data points, generally in the range of 500-1000s, as described in section 4.1.4. The sample was then left to relax

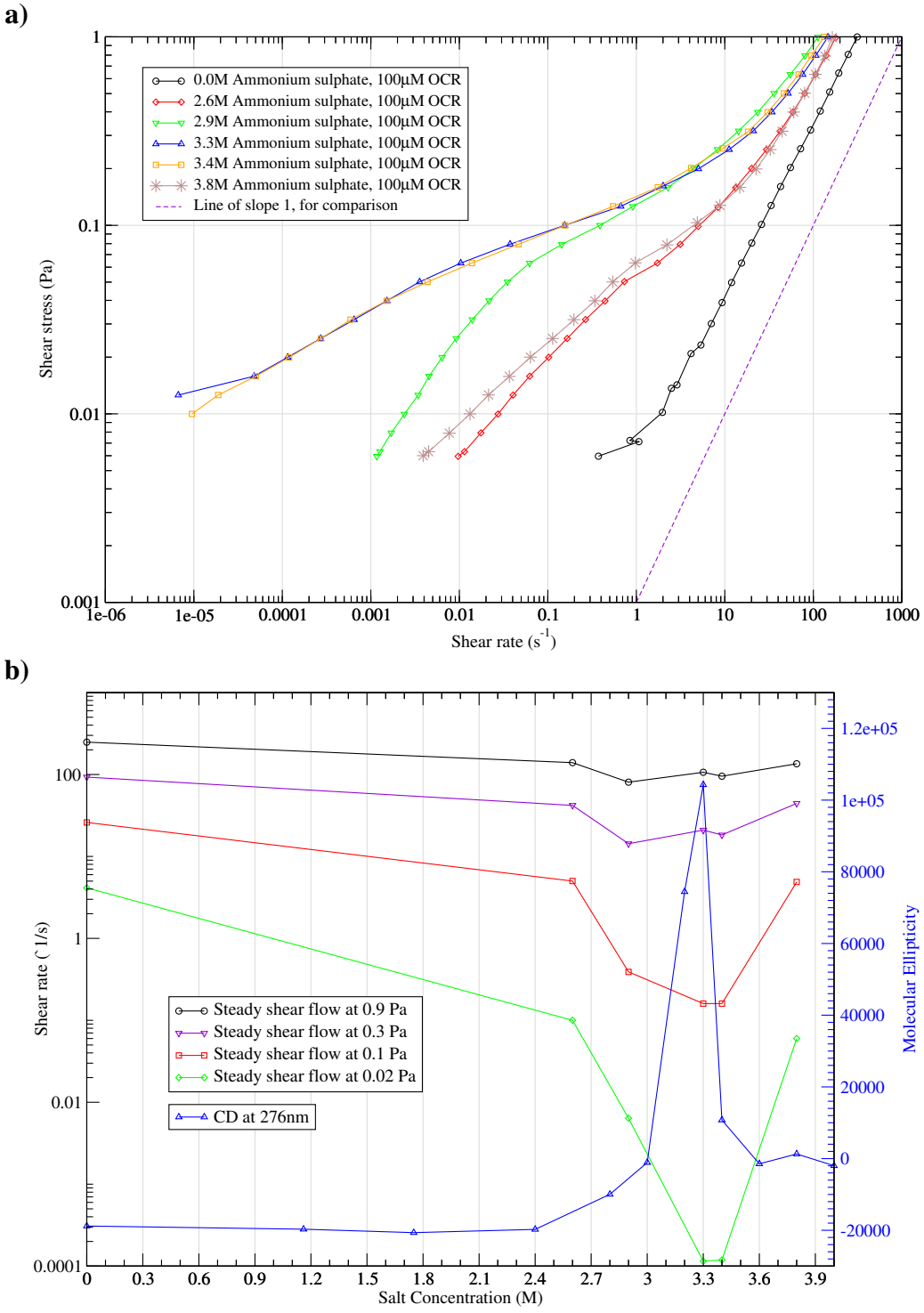


Figure 5.9: Steady shear rheology of 100μM ocr in a 20mM Tris-HCl, pH 8, buffer with varying concentrations of ammonium sulphate . Shear thinning is observed in the samples containing salt **(a)** along with a sudden decrease in shear rate around the gel point of the sample, which reflects circular dichroism results obtained from the same material **(b).** Within the gel region there appears to be a yield stress at around 0.2Pa.

for another 1000s and the strain in its final state measured to gauge its recovery (figure 5.10 **b)**.

A distinct yield stress at 0.2Pa was observed for numerous samples (an example of one such experiment is shown in figure 5.11**a)**. Initially the material resists the stress applied but at 0.2Pa it suddenly starts to flow. It becomes increasingly more inviscid with higher stresses before the underlying structure collapses above 0.5Pa. The fluid becomes superlinear as the stress overshoots ($B > 1$) before settling down to a steady-state, Newtonian flow [103]. Graph **b)** shows that the maximum strain recovered is between 40-50%. This is an order of magnitude higher than simple polymer systems and reflects the high elasticity of the protein gel. It decreases at higher stresses as the structure is overstretched.

5.3 Circular dichroism

5.3.1 Far UV

Far UV CD measurements of ocr were used to track changes in its secondary structure as the concentration of ammonium sulphate in the buffer was increased (figure 5.12a). These were compared to a sample containing 6M guanidine chloride, which completely denatures ocr [55]. Below 3.4M ammonium sulphate the CD spectra are comparable; above 3.3M they decrease in magnitude (by up to a factor of 3) and the secondary minimum at 222nm becomes less distinct (figure 5.12b). The unfolded sample in guanidine shows essentially no CD signal (six orders of magnitude change). The difference in ellipticity at 222nm between the current sample and the 3.2M sample, expressed as a percentage of the difference in ellipticity between the 3.2M sample and 6M guanidine sample, is shown in figure 5.13a)[2]. It indicates that while the secondary structure of ocr in the gel phase is unaffected by the salt, it may unfold by up to 50% above this.

The data was analysed using the DichroWeb program (section 4.2.1) to estimate for the ratio of components in the protein (α helix, β sheet and random coils). The result for ocr is shown in figure 5.13b). The analysis for β sheets is generally accepted to be less accurate than that for α helices; indeed we already know that there are no β sheets

[2]This can also be expressed as: $\frac{CD(3.2M)-CD(x)}{CD(3.2M)-CD(Gu)}$ evaluated at 222nm

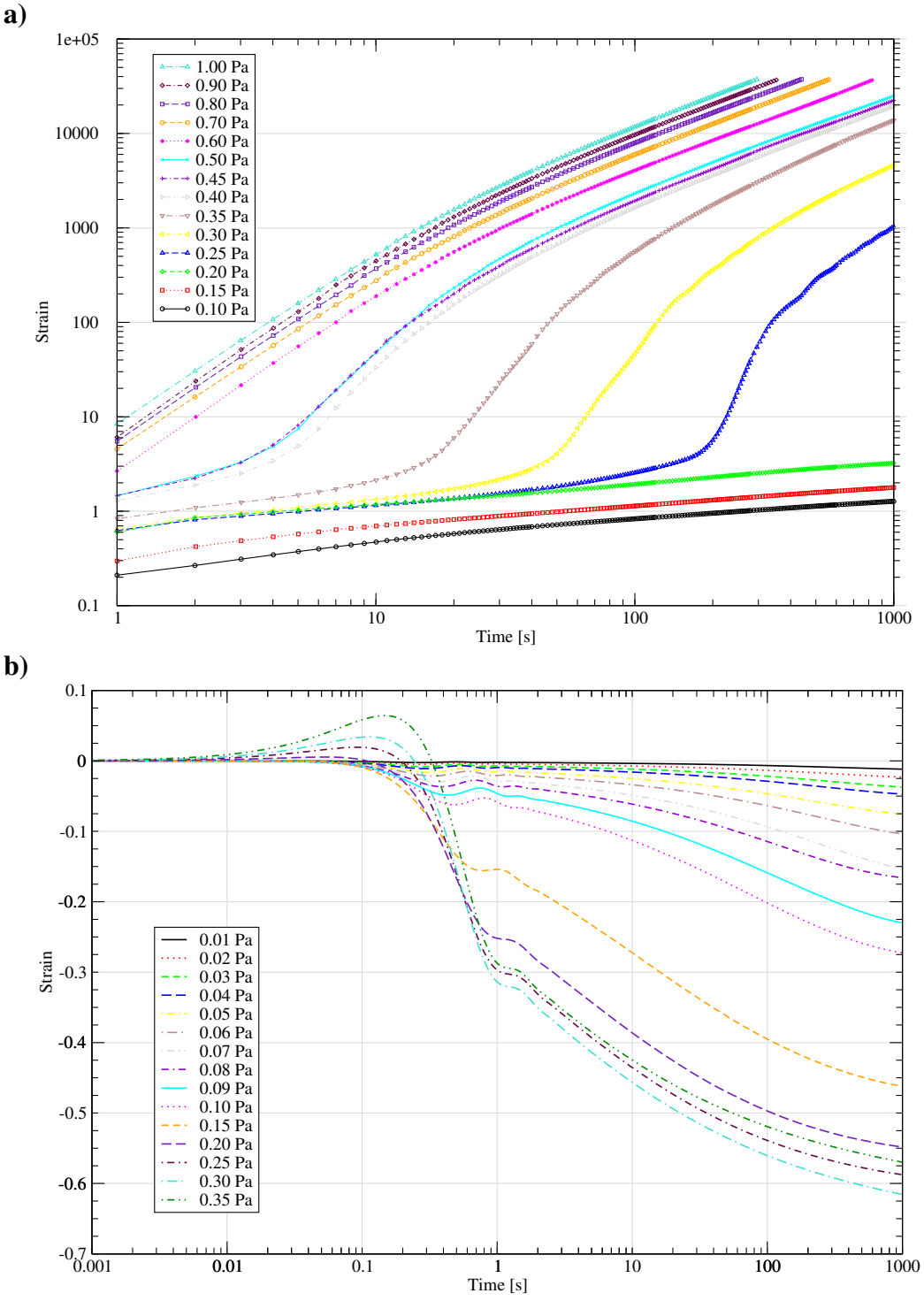


Figure 5.10: Data collected from creep **(a)** and recovery **(b)** experiments on a 36μM ocr gel. A range of stresses were applied from 0.01 to 1.0 Pa and the strain deformation measured. After 1000s the stress was stopped and the sample left for a further 1000s to recover. The strain was also measured during this time.

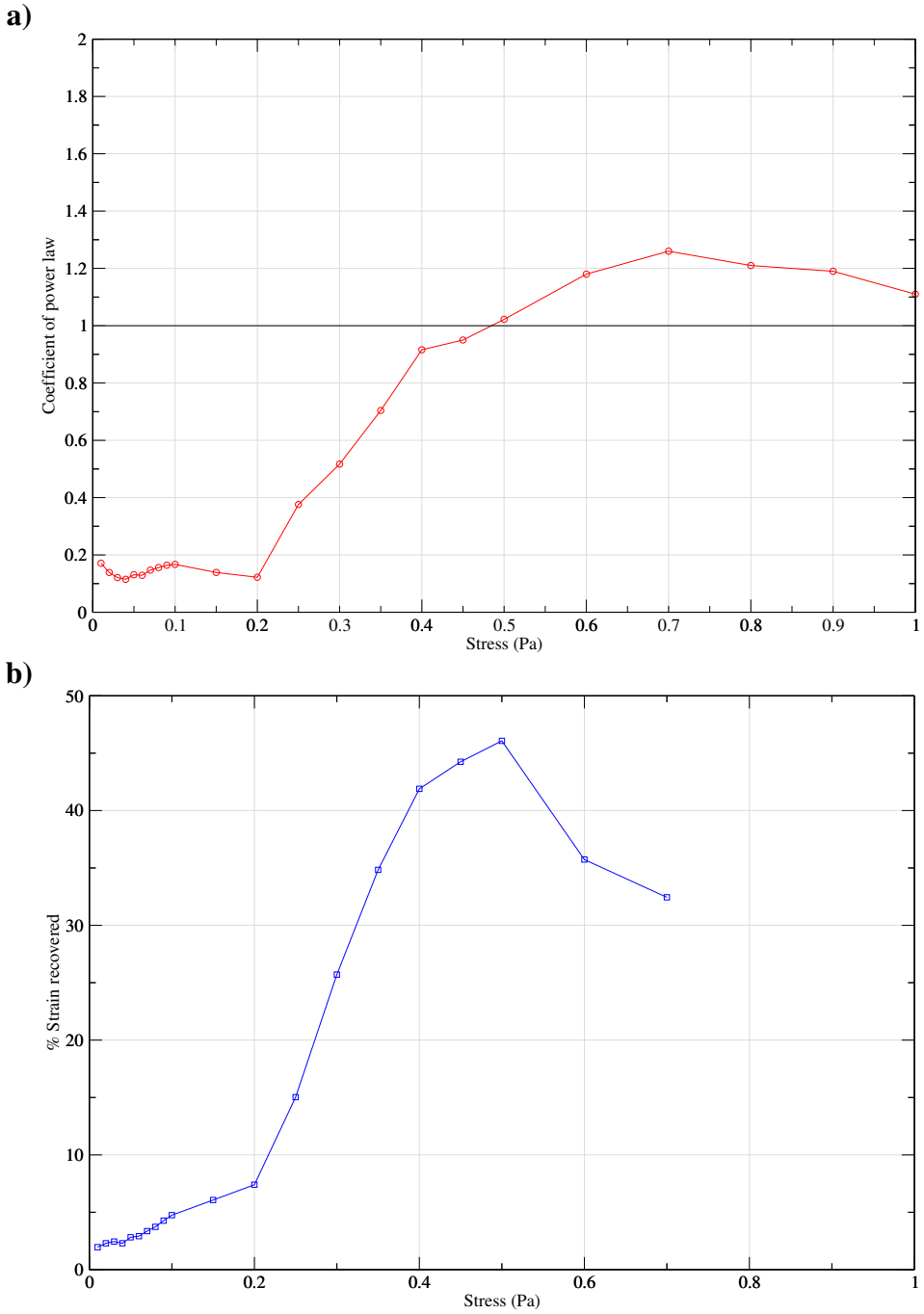


Figure 5.11: Analysis from creep **(a)** and recovery **(b)** experiments on a 100μM ocr and 3.3M ammonium sulphate solution in 20mM Tris-HCl, pH 8 buffer. The power law $y = Ax^B$ was fitted to the creep data (figure 5.10) and the power, B , plotted against stress **(a)**. A yield stress at 0.2Pa is observed and the maximum strain recovered once the stress has stopped **(b)** was 46%.

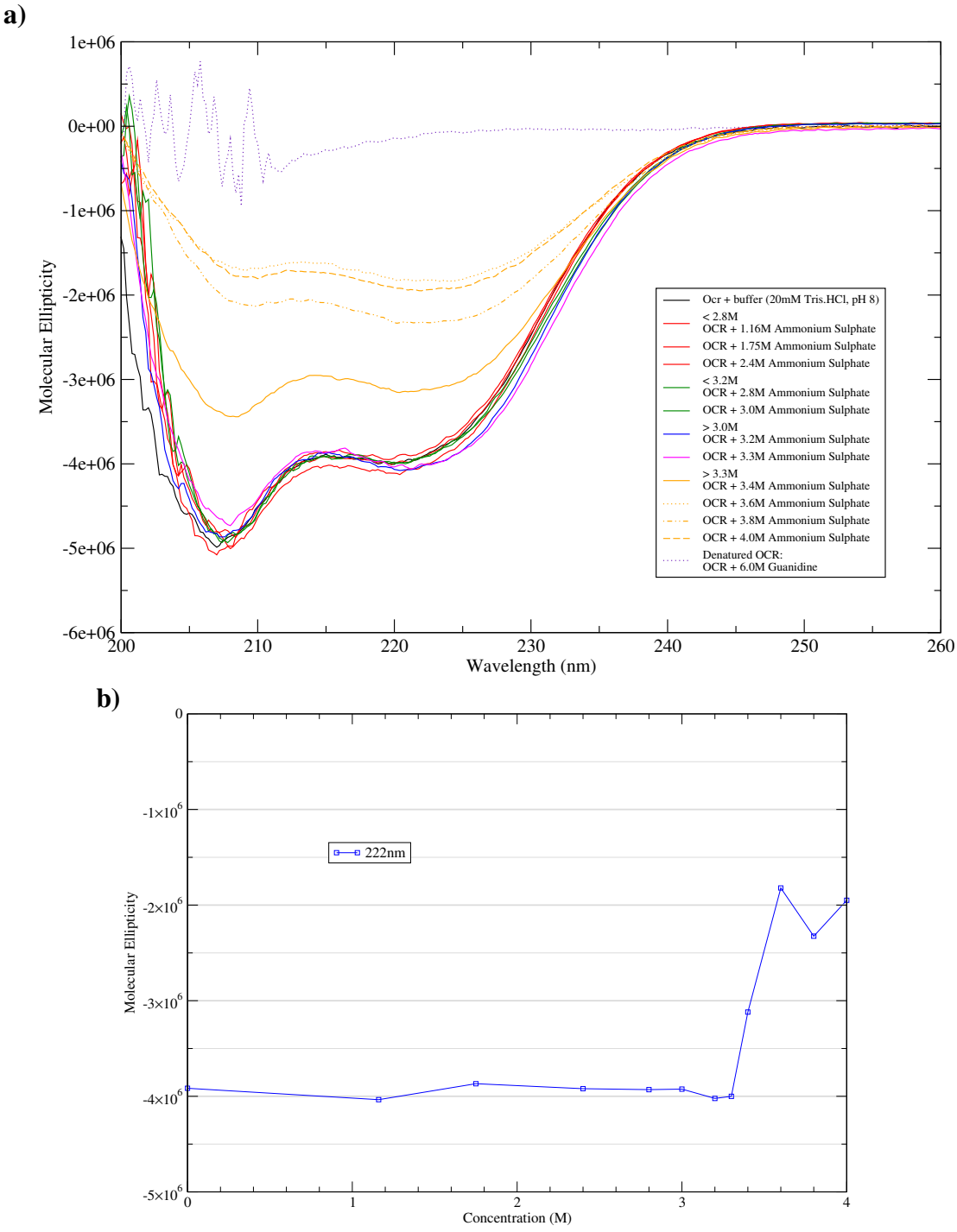


Figure 5.12: **a)** CD spectra at far UV wavelengths for a 30μM ocr solution with ammonium sulphate in varying concentrations. Above 3.3M the signals decrease in magnitude, by up a factor of 3. A sample prepared with 6M GuCl shows that the denatured protein gives no CD signal (i.e. a factor of 10^6 increase). **b)** Ellipticity at 222nm (from figure 5.12) over the range of ammonium sulphate concentrations used. Ammonium sulphate has no effect on ocr until 3.4M where the ellipticity increases by up to a factor of 2.

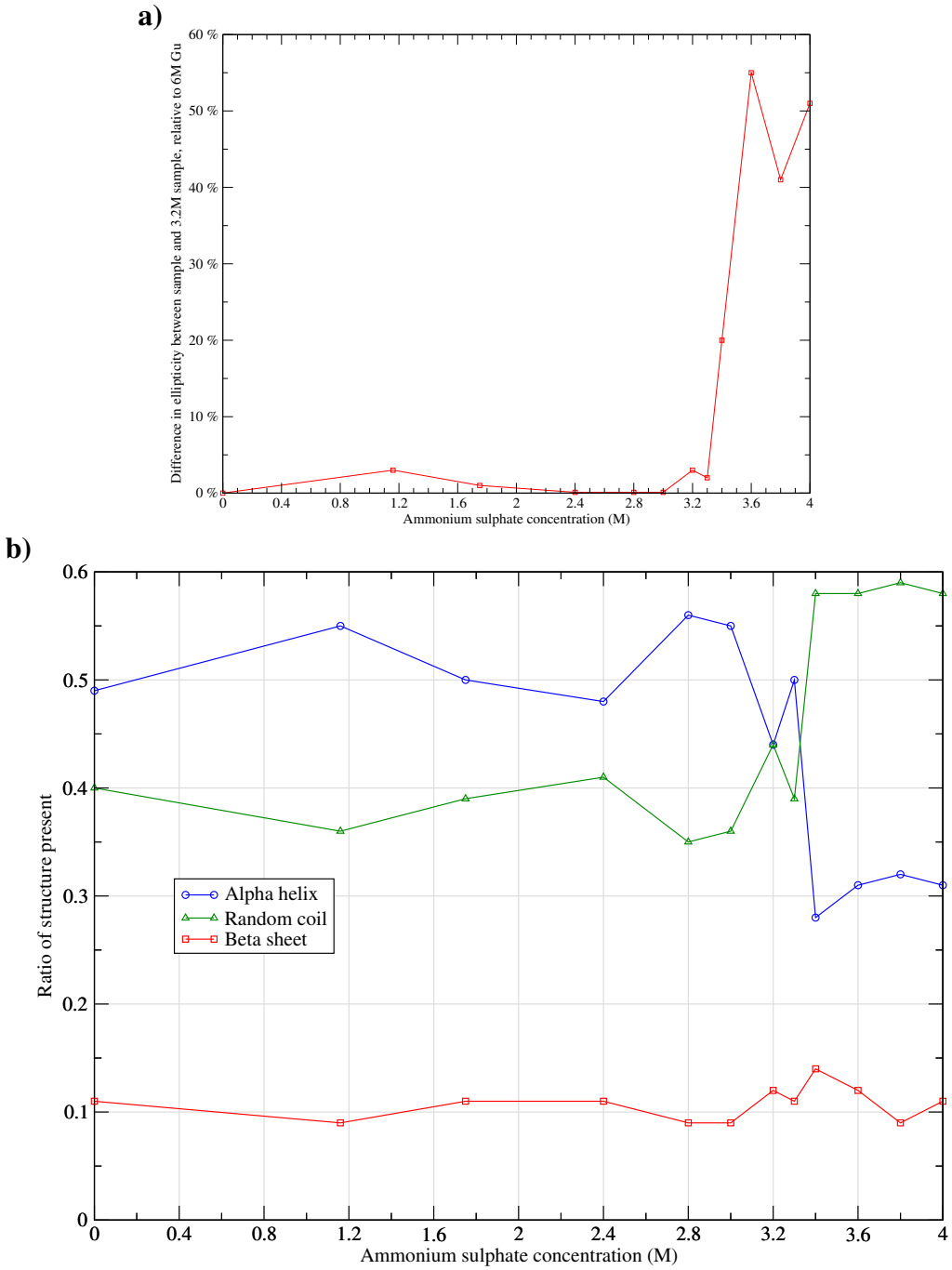


Figure 5.13: **a)** The difference in ellipticity at 222nm between the current sample and the 3.2M sample, expressed as a percentage of the difference in ellipticity between the 3.2M sample and 6M guanidine sample (fully denatured protein), as a function of the ammonium sulphate concentration. Ammonium sulphate has no effect on ocr until 3.4M where the protein begins to denature. **b)** Structural analysis of the far UV CD spectra showing a slight unfolding of alpha helices above the gel point. X-ray crystallography reveals no β-sheet content for ocr, so the constant value of 0.1 in the graph represents a systematic error in the calculations.

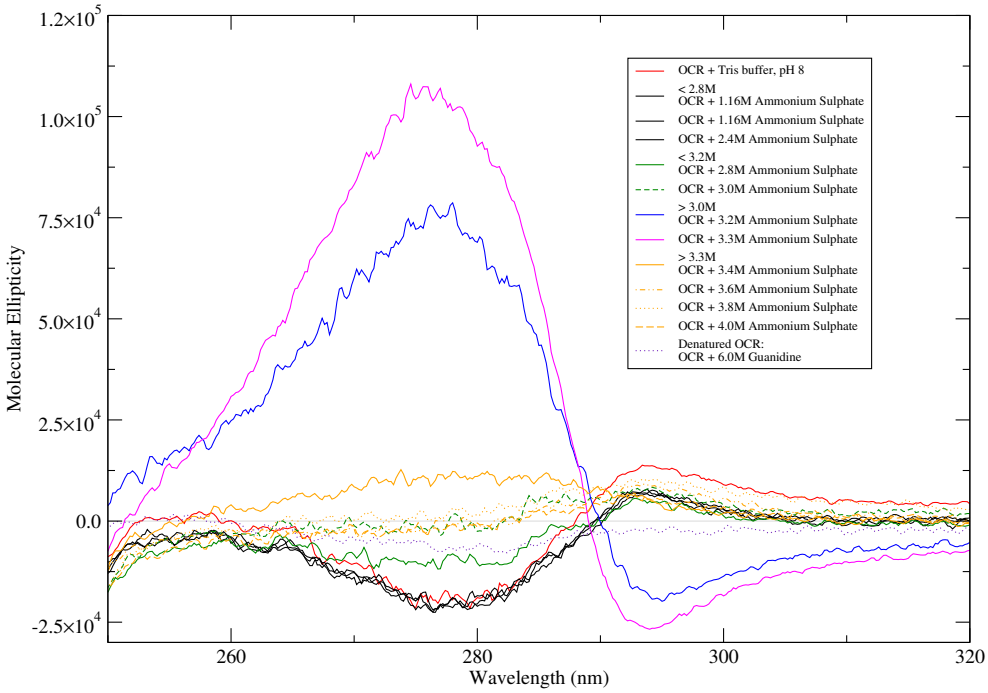
in ocr [66], contrary to the CD analysis at 0M ammonium sulphate. The β-sheet ratio of 0.1 for ocr, given by the CD analysis, remains constant over all ammonium sulphate concentrations. It can therefore be considered as a systematic error. One reason for the discrepancy is that the analysis is based on reference spectra of proteins with different structures and properties to ocr. The values obtained are therefore considered to be a first approximation only. Nevertheless, there is an apparent change in secondary structure at the gel point of the ocr-ammonium sulphate system. The α helix structure ratio changes by around 0.2 here, so even assuming an systematic error of ± 0.1, the change is noticeable. Although the quantitative analysis is inconclusive, it does imply a loss of α helices at salt concentrations above the gel point. In certain proteins the conversion of secondary structures to β-sheets has been shown to encourage the formation of amyloid fibrils [110]. However, it is clear from the denatured ocr sample, which gives essentially no CD signal, that there is no significant change induced by ammonium sulphate at the gel point.

5.3.2 Near UV

The experiments were repeated at near UV wavelengths (250 - 320 nm). Comparing CD spectra for a 30μM ocr solution under varying salt concentrations reveals differences in the tertiary structural environment (figure 5.14a). At 3.2M ammonium sulphate there is a dramatic increase from a small negative ellipticity to a strongly positive ellipticity. Samples with salt concentrations less than 2.8M have similar signals. Between 2.8 - 3.0M there appears to be a small transition before a sudden four fold jump at 3.2M. There is also a third region at concentrations of 3.4M and above where the signal drops just as sharply back down to zero. It does not return to the native 0M value though and visually corresponds to a less viscous gel. These observations are summarized in figure 5.14b), which gives the molecular ellipticity over the range of salt concentrations at three wavelengths: 294nm, 276nm and 220nm (inset). The striking appearance of a gel phase at 3.2M ammonium sulphate is clearly illustrated along with its weakening above 3.4M. The far UV CD also shows an increase in ellipticity at 3.2M but the change in magnitude is significantly lower than the near UV results.

The most common interpretation of near UV spectra is in terms of chirality. An alteration in the tertiary folding can place the aromatic side chains in chiral environments. The data in figure 5.14 shows an increase in left circularly polarized radiation,

a)



b)

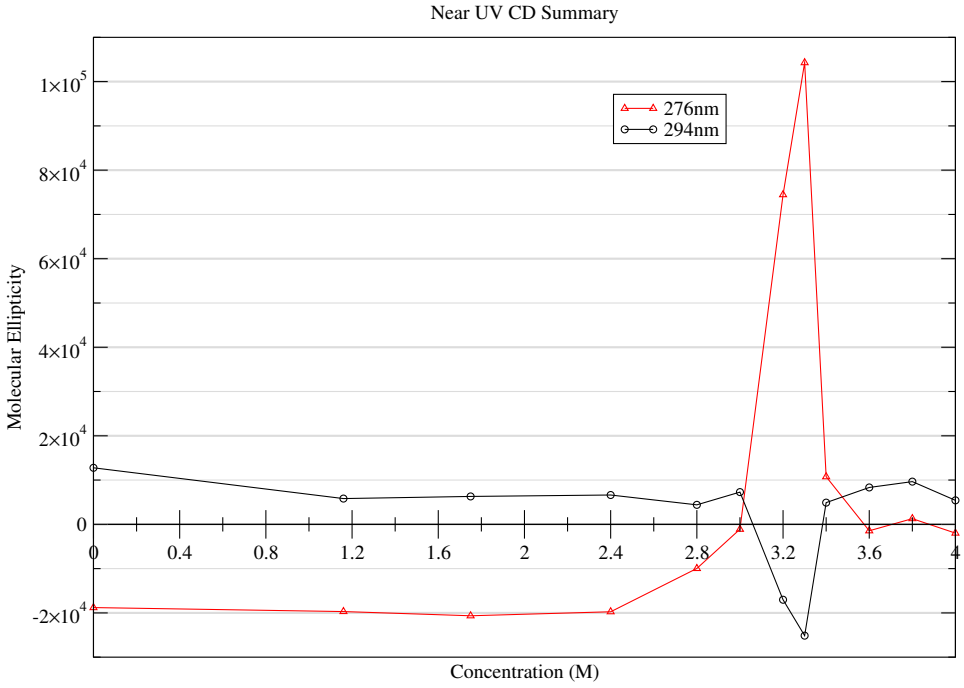


Figure 5.14: **a)** CD spectra at near UV wavelengths for a 30μM ocr solution with ammonium sulphate in varying concentrations. There is a sudden increase in signal at 3.2M ammonium sulphate then a decrease after 3.3M but not to the original values. **b)** a summary of how the CD signals changed over increasing salt concentration for selected frequencies (inset for far UV, main graph for near UV).

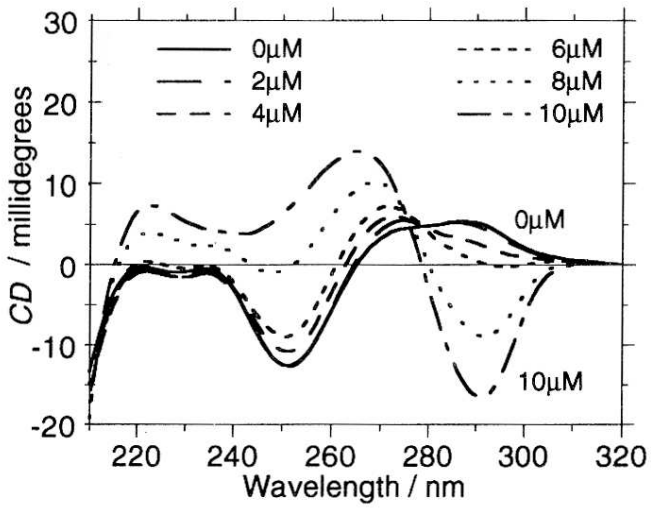


Figure 5.15: Near UV circular dichroism spectrum of 40μM B-form DNA in 20mM NaCl as a function of increasing spermine concentration. Above 7μM spermine the DNA changes from right-handed B-form to left-handed Z-form, accompanied by an inverse in the CD spectra. Diagram reproduced from [109].

possibly by the formation of left-handed helical structures. A similar reversal of CD signal can be observed with B-DNA. As spermine is added there is a sudden inversion in the CD spectrum (figure 5.15) [109]. The B-DNA, which has a right-handed twist, has changed state to Z-DNA, which has a left-handed twist.

5.3.3 Alternative salts

Proposing a model for the effect of ammonium sulphate on the ocr protein is compounded by the results shown in figure 5.16. Samples of ocr were made with seven other salts (with a variety of mono, di and trivalent cations and anions). Salt concentrations of 3.3M and 6.6M (or to saturation) were used to compare similar ionic strengths (table 5.2). Only ammonium sulphate formed a gel with ocr and it is clear from the near UV results that none of the other salts has any effect on its circular dichroism. In the far UV spectrum the signals retain their characteristic α helical minima at 222 and 208nm and the denatured sample (6M GuCl) shows that no salt causes major structural changes to the protein. The far UV data was analysed, as before, to obtain the secondary structure ratios (figure 5.17). The salts are grouped in terms of their ionic strength as $\frac{1}{3}I$, $\frac{2}{3}I$ and I, where I is the ionic strength of 3.3M ammonium

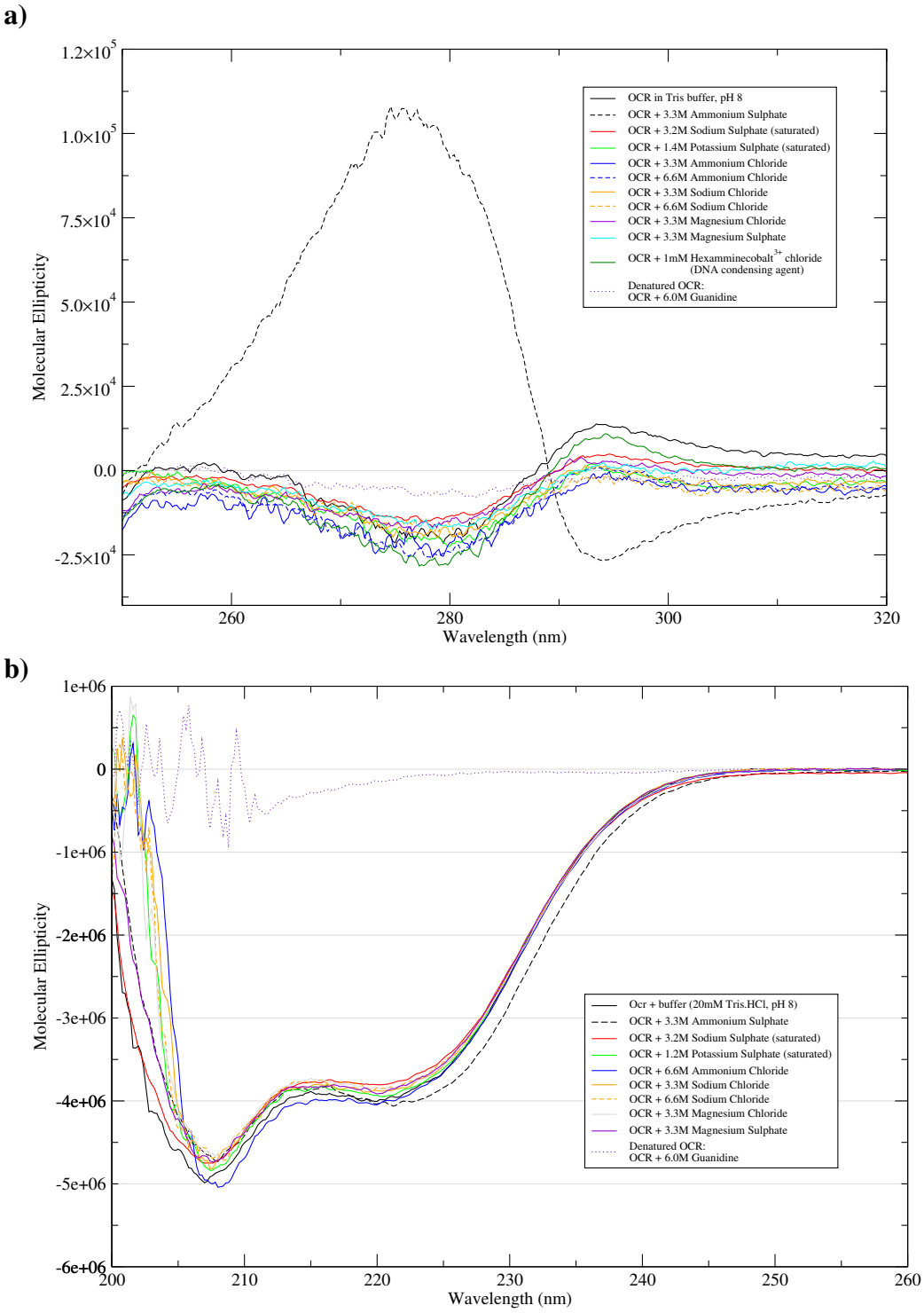


Figure 5.16: CD spectra at near **(a)** and far **(b)** UV wavelengths for a 30μM ocr solution with a range of salts. No other salt, whatever the valency, produces the same effect as ammonium sulphate. A sample prepared with 6M GuCl shows that the denatured protein gives no CD signal, while the other solutions do not change the conformation of ocr significantly.

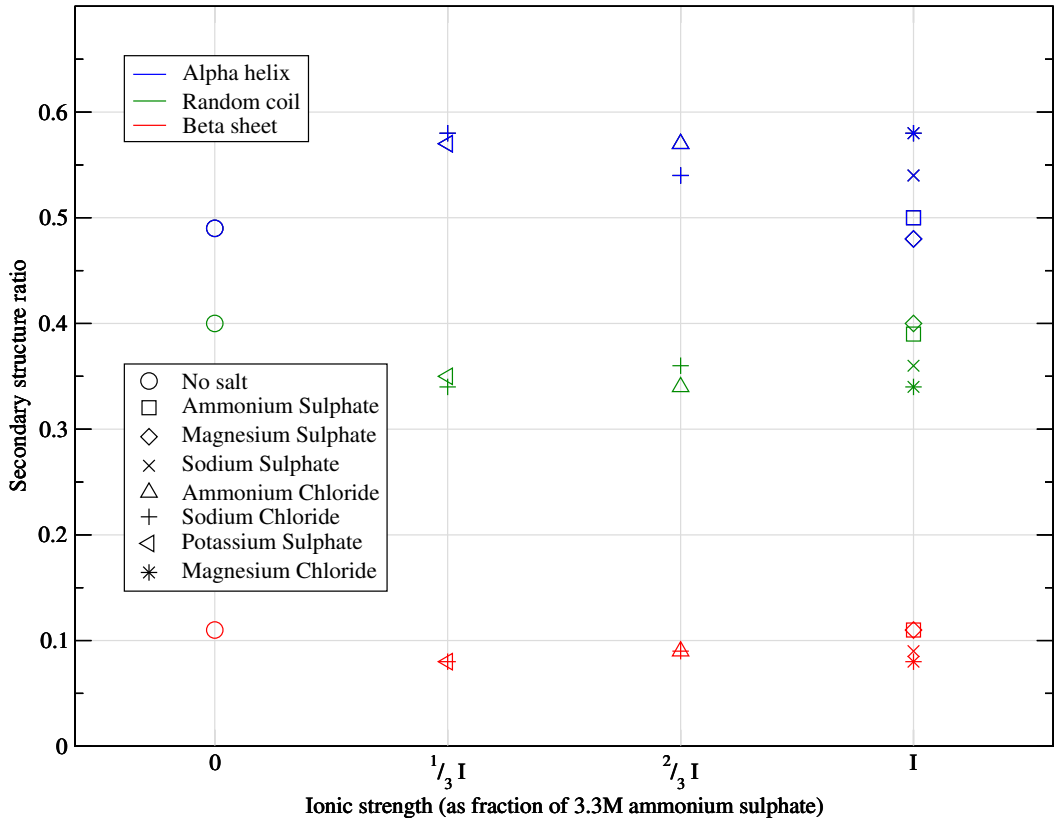


Figure 5.17: Structural analysis of the far UV CD spectra (figure 5.16b) for solutions of ocr in a range of salts. The data is categorised in terms of the salt ionic strength, as a fraction of 3.3M ammonium sulphate. The blue symbols represent alpha helix fractions, the green random coils and the red beta sheets. The secondary structure of ocr is not significantly affected by any salt.

Salt	Formula	Conc.	Cations ($\mathbf{L}^{-1}$)	cations : charges	$I = \frac{1}{2}\sum_i c_i z_i^2$
$X_2^+Y^{2-}$					
Ammonium Sulphate	$(NH_4^+)_2SO_4^{2-}$	0.0M	0	0	0
		3.3M	3.96×10^{24}	3.3×10^3	9.9
Sodium Sulphate	$Na_2^+SO_4^{2-}$	3.2M†	3.84×10^{24}	3.2×10^3	9.6
Potasium Sulphate	$K_2^+SO_4^{2-}$	1.4M†	1.68×10^{24}	1.4×10^3	4.2
X^+Y^-					
Ammonium Chloride	$NH_4^+Cl^-$	3.3M	1.98×10^{24}	1.6×10^3	3.3
		6.6M†	3.96×10^{24}	3.3×10^3	6.6
Sodium Chloride	Na^+Cl^-	3.3M	1.98×10^{24}	1.6×10^3	3.3
		6.6M†	3.54×10^{24}	2.9×10^3	6.6
$X^{2+}Y_2^-$					
Magnesium Chloride	$Mg^{2+}Cl_2^-$	3.3M	1.98×10^{24}	1.6×10^3	9.9
$X^{2+}Y^{2-}$					
Magnesium Sulphate	$Mg^{2+}SO_4^{2-}$	3.3M	1.98×10^{24}	1.6×10^3	13.2

Table 5.2: Comparing the valencies of the salts used in figure 5.16. At 30μM ocr has 1.2×10^{21} negative charges per litre. The 5th column shows the ratio of salt cations to negative charges on an ocr dimer. They are all of the order 10^3. († salt saturated)

sulphate. No salt has a significant effect on the secondary structure of ocr.

The trivalent salt hexamminecobalt is used to condense DNA. At 0.5mM concentrations it has been shown to induce a B-form to Ψ-form transition in calf thymus DNA [134]. The final, compact state being characteristic of aggregated DNA. Furthermore a 2-4 times increase in CD signal was observed at 270-280nm. However, no such effect was observed with ocr, despite its electrostatic resemblance to DNA.

5.4 Precipitation and pH

Section 2.6 discussed the intimate relation between protein charge and solution pH. The charge on ocr is most likely crucial to its ability to resist precipitation in salt solutions. The behaviour of an ocr gel at different pH values will reveal if precipitation occurs when the protein charge is perturbed enough and will provide an estimate to

this critical charge.

The following buffers were made using the procedures described in the Calbiochem guide to buffers [69]: 70mM citrate-phosphate (pH 1-7), 50mM acetate buffer (pH 3.6-5.6), 50mM hydrochloric acid-potassium chloride buffer (pH 1.0-2.6). Citrate-phosphate was used over the full buffer range (pH 1 - 8) to show the behaviour of ocr at different pH values is independent of buffer choice. The results were also corroborated using acetate buffer for pH 3 - 5, HCl-KCl for pH 1-2 and 20mM Tris-HCl for pH 8. Ammonium sulphate was added to each solution to a concentration of 3.25M and the pH value for each salt solution was measured to confirm it had not changed. Buffers were filtered through 0.5μm filters to remove any dust or large contaminants present.

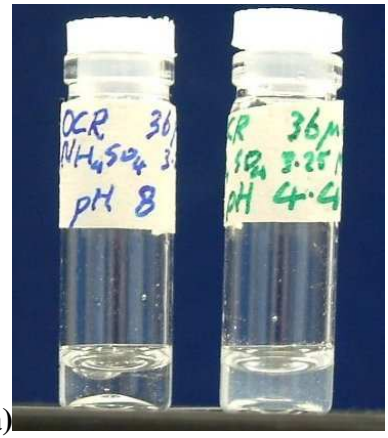
36μM solutions of ocr were made at a range of pH values using the buffered salt solutions. Above pH 6 the sample was transparent and formed a gel. Below pH 5 ocr precipitated, the solution became cloudy and remained a fluid. Figure 5.18 shows a solution at pH 8 and at pH 4.4, where it has precipitated and forms aggregates that can be seen under a microscope. Samples with pH values between 5 and 6 were made to probe this transition point. UV spectra were taken for each and can be seen in figure 5.19 a). The data for pH 6 and 8 show the characteristic shape for the electronic absorption by a protein. As the pH decreases there is a greater contribution from Rayleigh scattering from the globules of precipitated protein. For the best reproducibility each sample was made only after the previous sample's UV spectrum had been measured as there is a time-dependence for globule formation.

Each sample was then spun for 5 minutes at 18000g and 4°C in a centrifuge. The pH 8 and 6 samples remain stable gels and no precipitate was observed. The supernatants were transparent for all the samples with increasing pellet sizes at lower pH values. The UV absorbance for each supernatant was measured and can be seen in figure 5.19 b). A gradual transition is revealed from no precipitate at pH 8, which has an identical spectrum to the unspun one, to almost completely precipitated at pH 4.

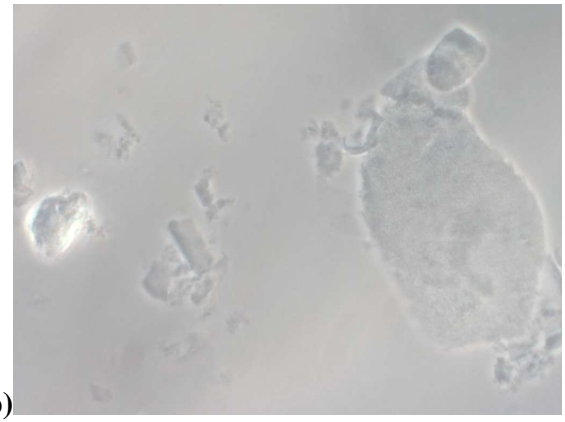
Rayleigh scattering is inversely proportional to λ^4. Since the spectra below pH 8 is a combination of scattering and absorbance at pH 8, A_{pH8}, they can be fitted to:

$$\frac{a}{\lambda^4} + bA_{pH8} + c \tag{5.1}$$

where a, b and c are the parameters to be calculated. The best-fit curves can be seen in figure 5.20 a) and their parameters in table 5.3. The Rayleigh scattering co-



a)



b)

Figure 5.18: a) Solutions of ocr in 3.3M ammonium sulphate at pH 8 and pH 4.4, where the protein has precipitated, leaving a cloudy solution. b) Microscope image of aggregates in the pH 4.4 solution. A plan 40×, 0.65 Ph 2 objective was used. The image is 200μm wide.

efficient (a) and proportion of absorbance at pH 8 (b) are plotted against pH in figure 5.20 b). This shows the gradual transition from transparent gel above pH 6 to cloudy, precipitated solution below pH 5. The cross over point occurs around pH 5.6, which corresponds to a charge in the order of -48e (according to the basic titration curve in figure 2.17).

Images of ocr solutions with pH values spanning the precipitation boundary were obtained using atomic force microscopy (section 5.5.5). These support the results from UV spectroscopy that when there is insufficient charge on the ocr proteins, they are unable to form a network or a gel. The link between pH, charge and precipitation is further discussed in section 7.3.

5.5 AFM

Under an AFM, fibrils were instantly and clearly observed for a 36μM ocr and 3.2M ammonium sulphate gel (20mM Tris-HCl, pH 8). Figure 5.21 shows a network of distinct fibrils, spanning the gel. Fibril dimensions will be discussed below but were on average 2.6-10.4nm in height and ~1μm long. AFM provides a unique opportunity to probe the conditions under which ocr fibrils, and gels, form. A range of samples

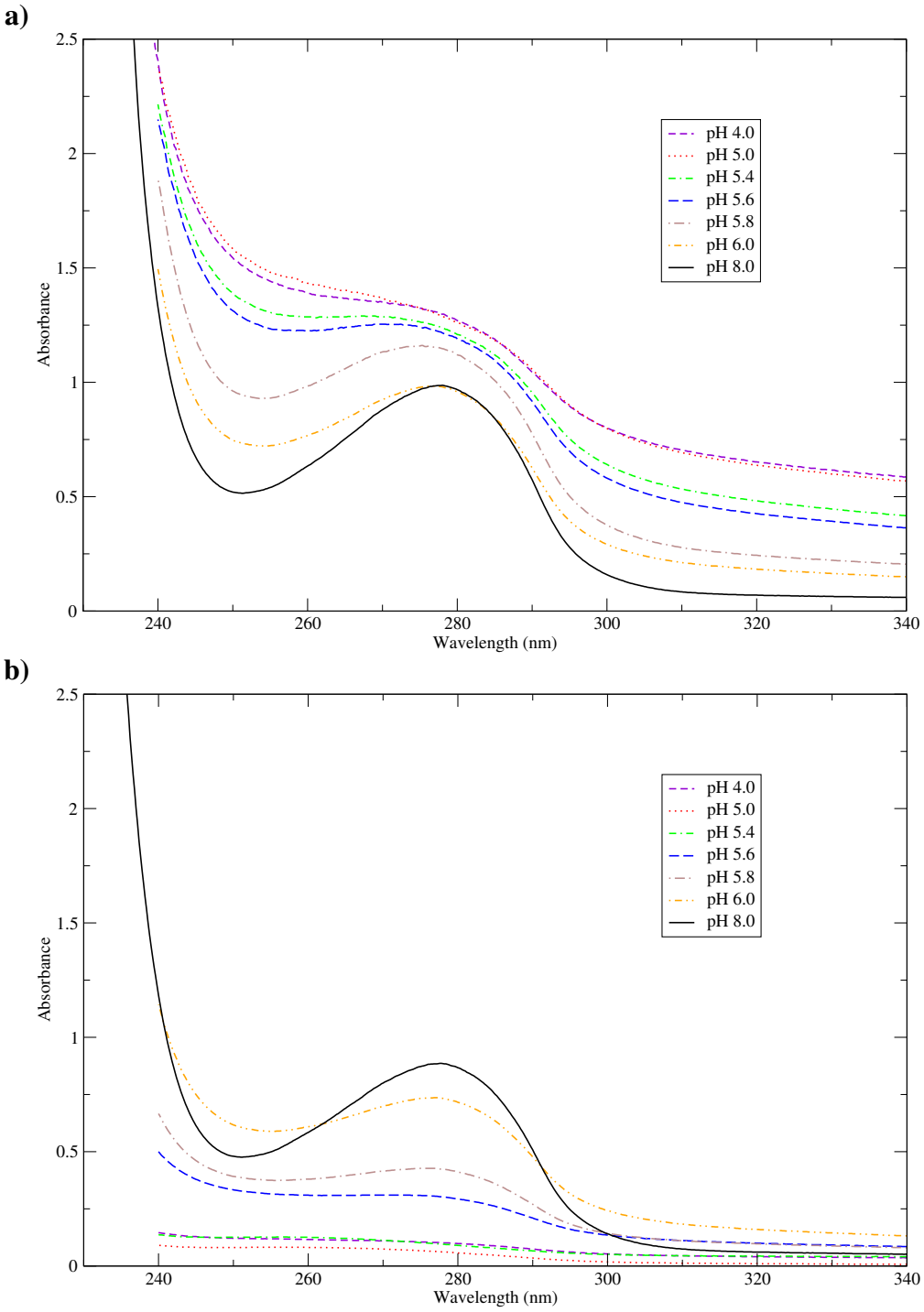


Figure 5.19: **a)** UV spectra of a 36μM ocr sample in buffered salt solution (3.25M ammonium sulphate) at pH 4 to 8. The transition between a transparent, gel at pH 8 and cloudy, precipitated protein solution at pH 4 is revealed by the increasing Rayleigh scattering at lower pH values. **b)** UV spectra of the supernatant from each sample after 5 minutes in a centrifuge. The pH 8 sample is unchanged. As the pH decreases more protein precipitates and is pelleted in the centrifuge. This is reflected in the decreasing absorbance of the supernatants.

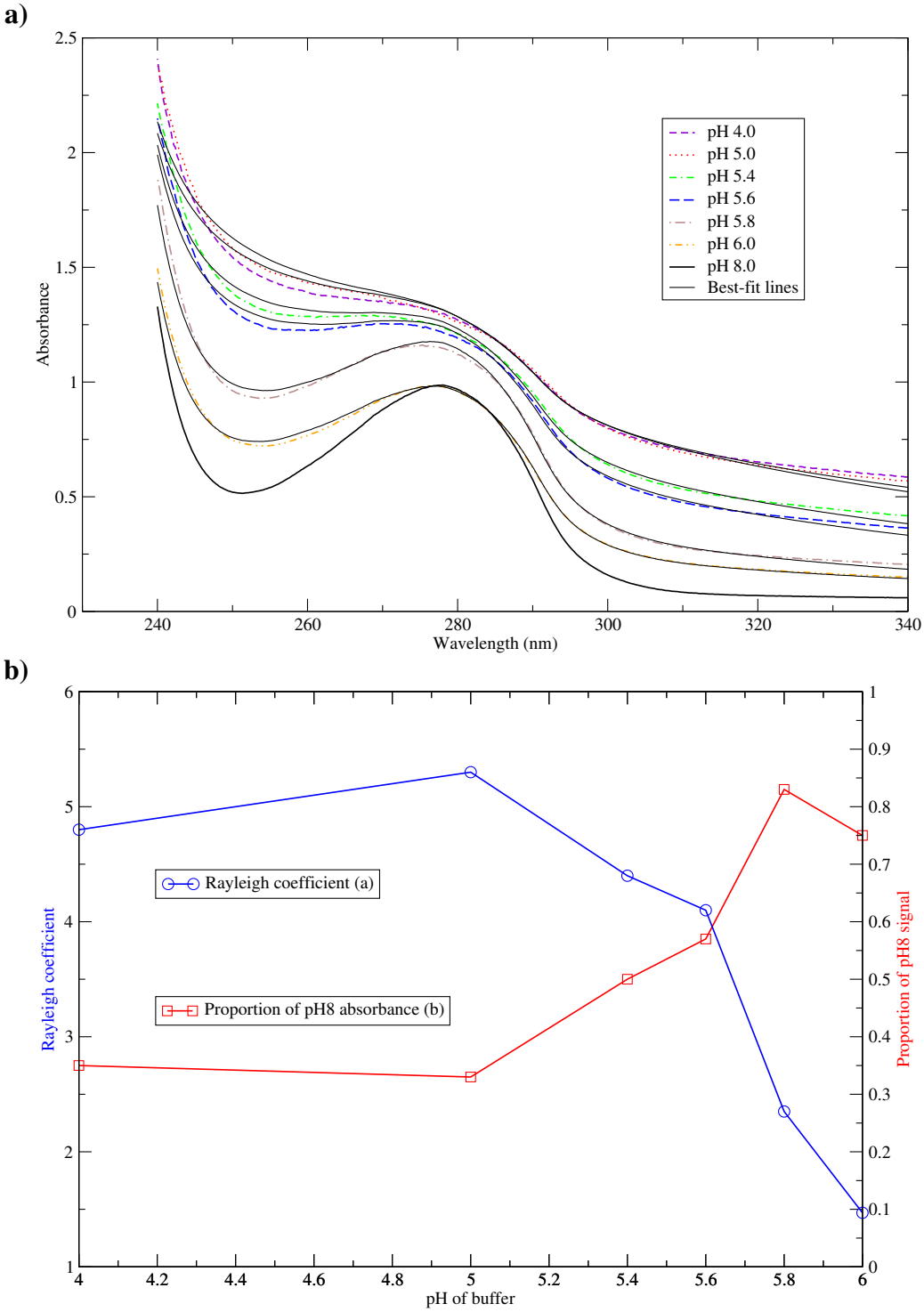


Figure 5.20: **a)** UV spectra of ocr/salt solutions at different pH values (figure 5.19 a) with best-fit lines based on equation (5.1) superimposed. **b)** The parameters used in the best-fit curves against pH, showing the gradual transition from transparent gel above pH 6 to precipitated solution below pH 5.

pH	a ($\times 10^9$)	b	c
4.0	4.8	0.35	0.16
5.0	5.3	0.33	0.11
5.4	4.4	0.50	0.02
5.6	4.1	0.57	-0.01
5.8	2.3	0.83	-0.04
6.0	1.5	0.75	-0.01

Table 5.3: Best-fit parameters to equation (5.1): $\frac{a}{\lambda^4} + bA_{pH8} + c$, combining Rayleigh scattering and absorption for 1mg/ml ocr in 3.25M ammonium sulphate buffered solutions of varying pH. Parameters are plotted in figure 5.20 b).

were prepared, covering different salt concentrations, protein concentrations and pH values. The effect of diluting a gel with buffer solution was also studied.

5.5.1 Salt concentration

On average three areas were probed per sample. For those with little obvious structure more searches were performed to ensure no details were missed. Where fibrils were abundant around 100 per sample were randomly chosen to measure heights and lengths. Care must be taken to choose a level area when measuring fibril heights. Fibrils that lie over crystals of salt or aggregates of proteins can appear larger than in reality. This also limits the number of fibrils that can be accurately measured and is most noticeable for solutions either side of the gel phase. Within the gel phase the fibrils are flat and distinct, resulting in a narrow spread of measured heights. A scatter chart comparing the number of fibrils measured and the fibril heights at each salt concentration is shown in figure 5.23 and discussed below.

Seven concentrations of ammonium sulphate were used with 36μM ocr samples (20mM Tris-HCl, pH 8). The AFM images obtained for each are shown in figure 5.22. At 2M ammonium sulphate there were no fibrils. After searching many different areas in vain a single, tenuous ridge was found in the otherwise rough, flat landscape. This may or may not be a fibril but it is clear that, as expected, ocr does not aggregate or form structures on a large scale at this salt concentration. Towards the gel point a small area of possible fibrils was found at 3.0M, again after much searching, in an increasingly

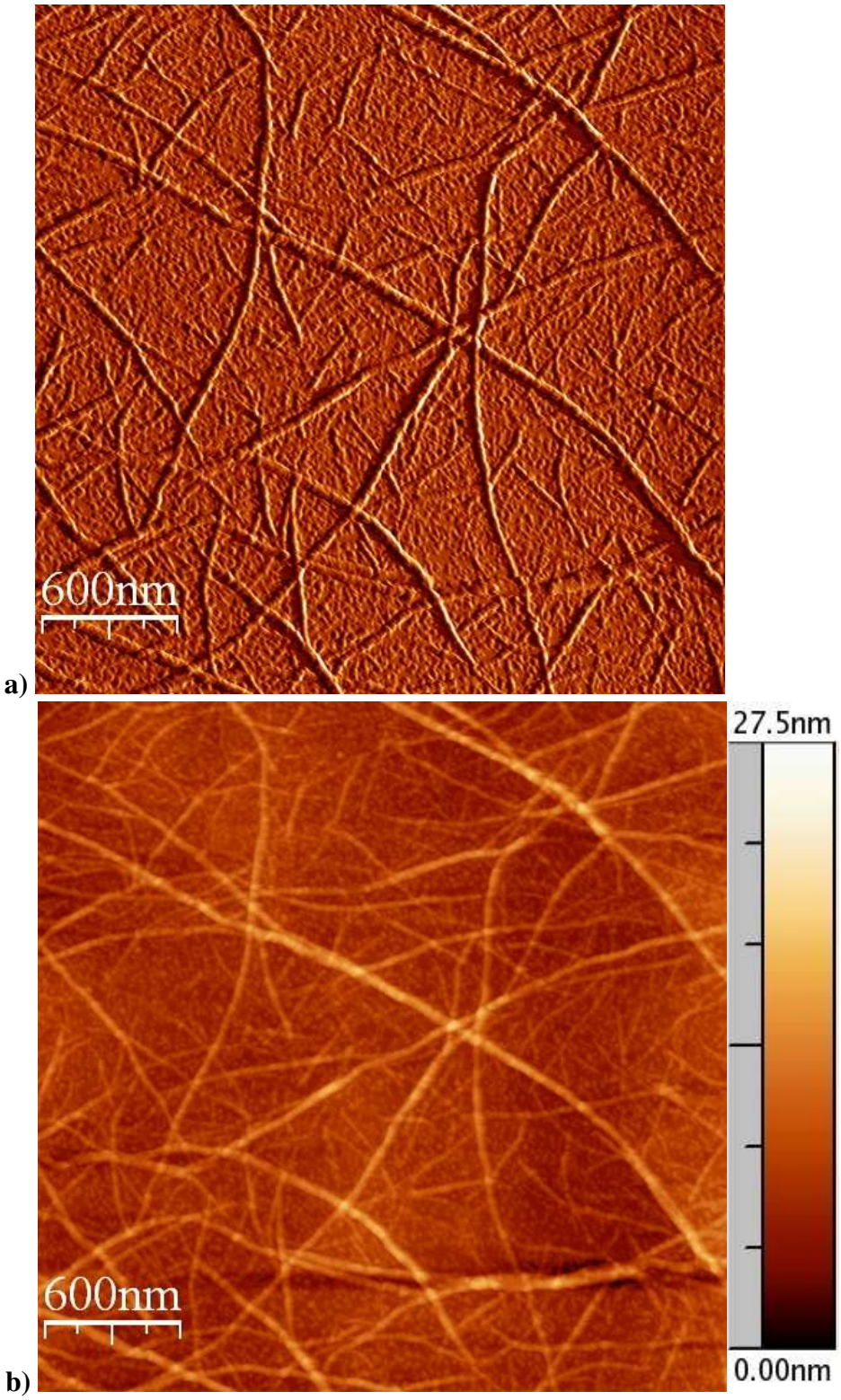


Figure 5.21: AFM images of a 36μM ocr sample in 20mM Tris-HCl, pH 8 with 3.2M ammonium sulphate. **a)** Amplitude data **b)** height data. Fibrils are clearly seen spanning the gel with heights of ~10nm and some with lengths over 1μm.

rough terrain. By 3.1M fibrils were suddenly, and easily, observed. They tend to cluster or bunch, making height and length measurements tricky. They also appear more flexible and do not form a clearly distributed network. Within the gel phase, at 3.2M and 3.3M, the fibrils are predominantly straight, well separated and span the sample. Above the gel phase, at 3.4M and 3.5M, the fibrils become less defined and begin to cluster again. Protein aggregates also begin to form as the network appears to be collapsing. Measurements of heights and lengths once again become harder to perform accurately.

Each fibril height measured is plotted as a function of ammonium sulphate concentration in figure 5.23 along with the total number of fibrils measured for each sample. The data collected for 3M is limited and inconclusive. The few fibril-like objects observed were indistinct and obscured by aggregates, making accurate measurements difficult. Fibrils were more developed for higher salt concentrations, improving the quality and quantity of measurements. Gelled samples, in particular, displayed the tightest distributions due to their clear spacing and lack of aggregates. Fibril heights varied between 2.6nm and 10.4nm, the thickness and length of the ocr dimer (modelled as a prolate ellipsoid, as described in figure 2.6). Outside of the gel phase the spread of data increases. The average fibril height also appears to increase. Both these effects are likely to be caused by the bunching of fibrils and the presence of aggregates, which makes defining a zero height reference difficult.

A histogram of this data is shown in figure 5.24. Multiples of 2.6nm were used to categorise the fibril heights as this is the thickness of an ocr protein. The number of fibrils in each height category is plotted as a fraction of the total number of fibrils measured at the respective salt concentration. The histogram is repeated in 5.24b) superimposed on a normal distribution for the fibril heights at each salt concentration. A Gaussian of the form

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{5.2}$$

was fitted to each sample's data , where μ is the mean and σ the standard deviation of the height data. The results for fibrils in the gel phase (3.2M, 3.3M) are clearly more defined than samples of higher or lower salt concentrations. They show that fibrils in the gel phase are on average 2 to 3 times thicker than a single ocr protein. Fibril heights vary between 3 to 7 times the thickness of an ocr protein outside the gel,

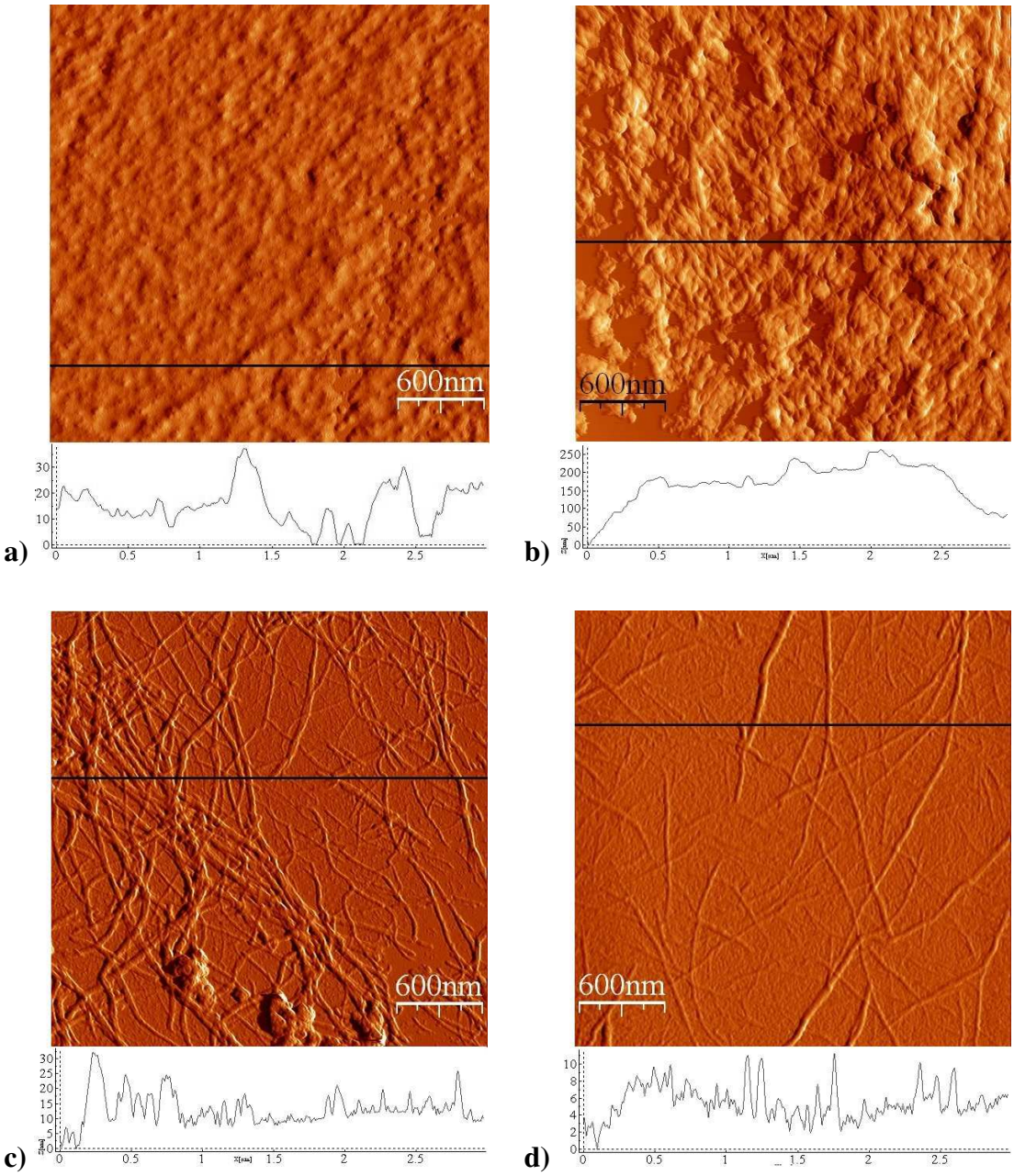
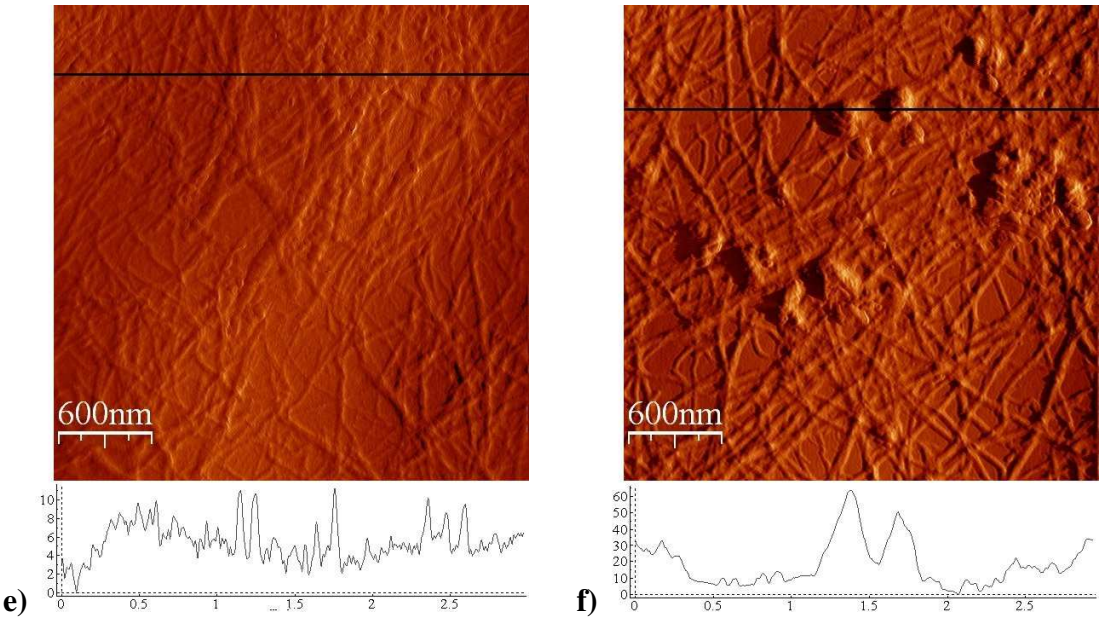


Figure 5.22: [continued overleaf] AFM images of 36μM ocr samples in 20mM Tris-HCl, pH 8 at salt concentrations spanning the gel transition: **a)** 2.0M, **b)** 3.0M, **c)** 3.1M, **d)** 3.3M (gel), **e)** 3.4M and **f)** 3.5M ammonium sulphate. At 2M only a single, tenuous fibre was found after much searching. As the salt concentration increases fibres become more distinct and begin to form a distributed network. Above the gel point the network becomes less well-spaced and aggregates begin to form. Line profiles beneath the images provide a measure of feature height (nm) along the scan line highlighted in each image. [figure continued overleaf]



[Figure 5.22, continued from previous page] AFM images of 36μM ocr samples in 20mM Tris-HCl, ph 8 at salt concentrations spanning the gel transition: **a)** 2.0M, **b)** 3.0M, **c)** 3.1M, **d)** 3.3M (gel), **e)** 3.4M and **f)** 3.5M ammonium sulphate. At 2M only a single, tenuous fibre was found after much searching. As the salt concentration increases fibres become more distinct and begin to form a distributed network. Above the gel point the network becomes less well-spaced and aggregates begin to form. Line profiles beneath the images provide a measure of feature height (nm) along the scan line highlighted in each image.

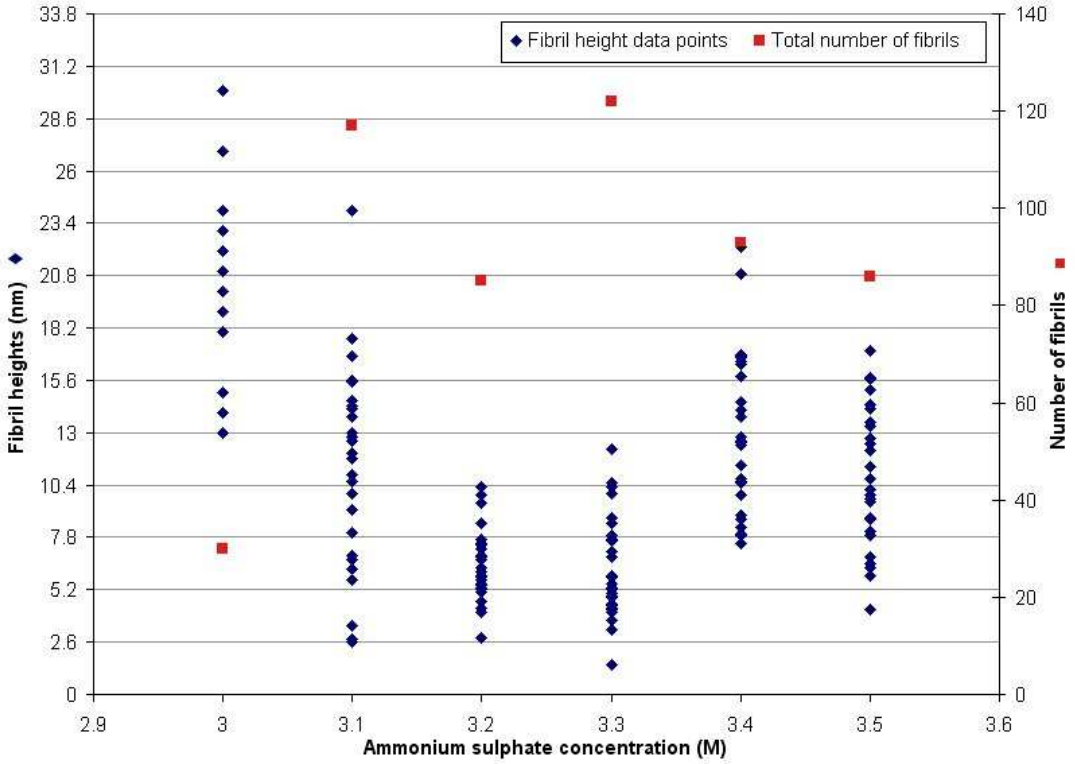


Figure 5.23: Scatter chart showing all of the fibril heights measured (♦), and the total number of fibrils (■), at each salt concentration. The distribution is tightest for 3.2M and 3.3M, inside the gel, where it is centred around 6.2nm. The fibril height scale is marked in units of 2.6nm, the thickness of an ocr protein.

as indicated by the scatter chart. However, this may be an artifact of the measurement due to aggregates and uneven sample surface, rather than an inherent increase in fibril height.

Results for fibril lengths are more in agreement over the range of salt concentrations. A histogram of the data is shown in figure 5.25. The average fibril length across the salt concentrations, both in solution and gel phase, is around 1μm. When fibrils bunched together it was difficult to determine the start and end of a fibril. Measurements were also limited to a 3μm square view but it was uncommon to find fibrils that spanned the whole image.

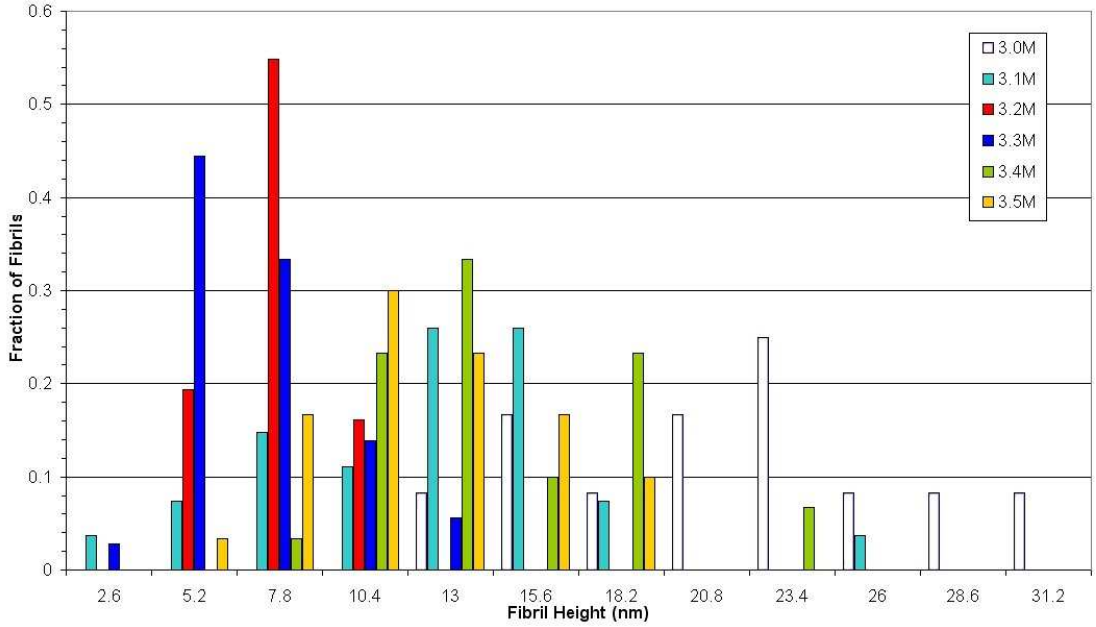
Measuring fibril widths, in the plane of the AFM image, is not as accurate as measuring their height due to the relatively large size of the tip, as described in section 4.3. The widths of fibrils in the gel predominantly varied between 10-50nm. Equation 4.13 takes account of the tip convolution and estimates the width at half height of the fibrils to be $2\sqrt{RD+\frac{D^2}{4}}$, for a tip of radius R and fibril of diameter D. If we assume ocr fibrils are made up of a single chain of dimers, then their range of widths would be 2.6-10.4nm. A tip of radius 10nm would therefore result in observed widths at half height of 10.5-22.9nm. These are an order of magnitude higher and are comparable to the measured images. It is also possible that multiple chains twist round each other to form larger fibrils, but when drying on the AFM substrate they collapse slightly, leaving fibrils with larger widths than heights.

5.5.2 Repetitions and junctions

Two particularly interesting features were observed in the above samples. Firstly, a number of fibrils in the gel phase appear to have a regular structure along their length, as shown in figure 5.26. Bands 30nm wide are visible, aligned almost perpendicularly to the fibril axis. In fibrils of different orientations the bands remain perpendicular so are not an artifact of measurement. The distance between these features was measured from the left edge of one band to the left edge of the next, rather than measuring between the centres as they are hard to define. The variation in height along the top of the fibril is shown in 5.26c). There is no regular height difference between peaks; the drops varying between 1 - 2.5nm. This could be due to uneven drying of fibrils or an uneven surface beneath the fibril (for instance, other fibrils or salt deposits).

Secondly, in some images there appear to be junctions or branches in the fibrils, as

a)



b)

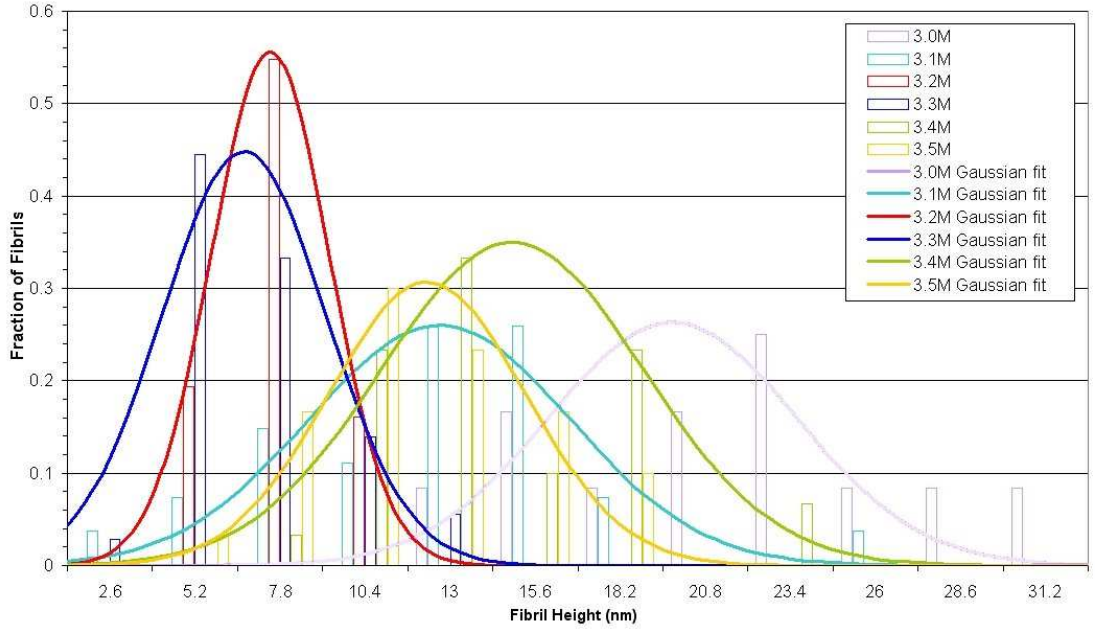


Figure 5.24: a) Histogram of the fibril heights as a fraction of the total number of fibrils measured for the respective salt concentration. Height categories are multiples of 2.6nm, the thickness of an ocr protein. b) Gaussian distribution of the height data for each salt concentration (superimposed on histogram a). The distribution in the gel phase is tighter than outside it, and gives an average height of 2-3 times the thickness of a single ocr protein.

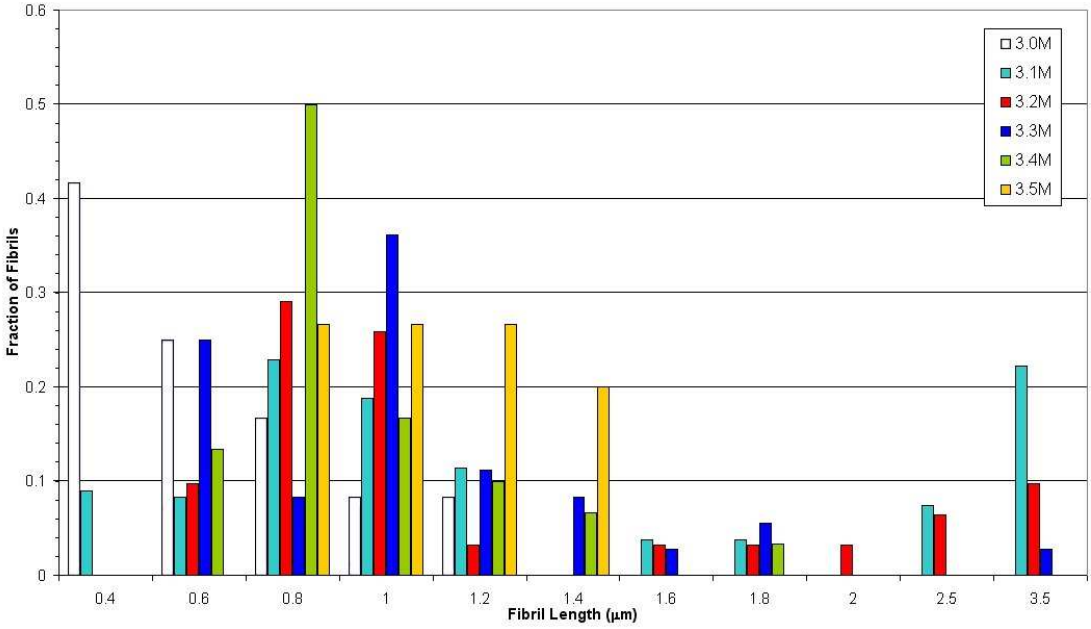


Figure 5.25: Histogram of fibril lengths as a fraction of the total number of fibrils measured for the respective salt concentration. The average fibril height was 1μm.

shown in figure 5.27. In some cases this is simply where two fibrils come in contact or overlap and appear merged in the image. However, there are instances where the fibril height at the juction is less than the combined height of two fibrils. For example, the juction in figure 5.27b) is 8.9nm in height, whereas the individual fibrils leaving the juction have heights of 5.9nm and 6.7nm. It is seen to occur with salt concentrations between 3.1M to 3.4M, so is not a feature of the gel alone, nor is it a feature of imperfectly developed fibrils. They are discussed further in section 7.4.

5.5.3 Protein concentration

To see how dependent fibril formation was on protein concentration, two samples of lower ocr concentration were prepared in 3.3M ammonium sulphate. The results of 0.36μM and 3.6μM ocr samples can be seen in figure 5.28. In both over 5 areas were searched. At 3.6μM ocr is still able to form fibril-like structures. They are concentrated in a small area, presumably because of the low numbers of proteins, and do not appear fully formed. At 0.36μM no fibrils or aggregates were observed; the protein concentration is apparently too low.

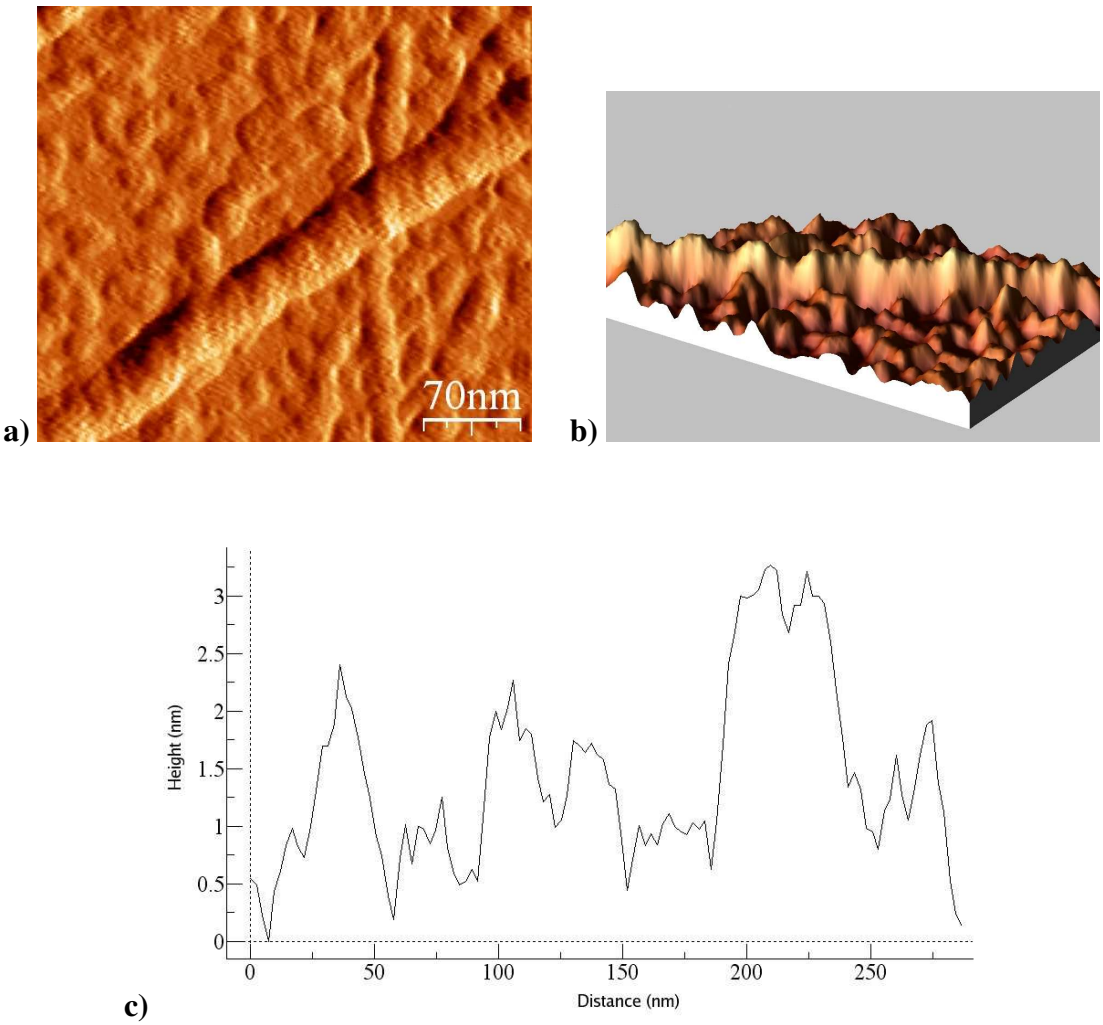


Figure 5.26: AFM image of an ocr fibril formed in 3.2M ammonium sulphate. Regular bands perpendicular to the fibril axis are seen, each with a width of 30nm. Image b) is a 3-dimensional representation of image a). c) Profile along fibril length showing the regularly spaced bands and a height variation of 1-2.5nm.

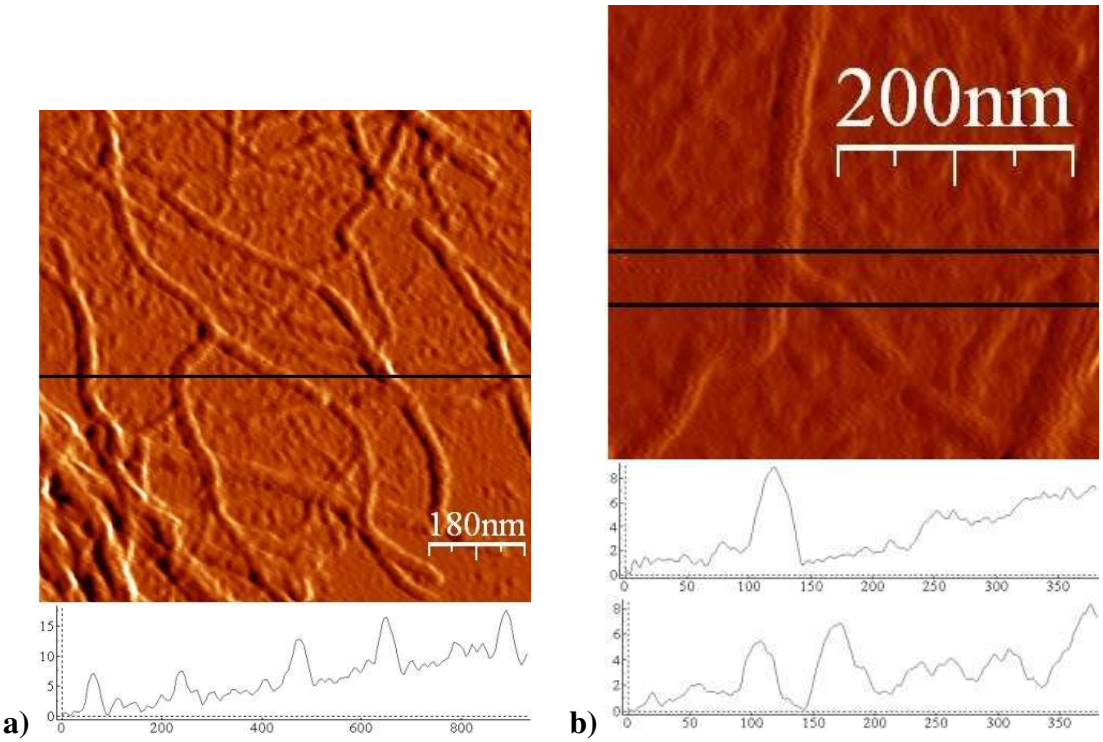


Figure 5.27: AFM images of ocr fibrils showing junctions between fibrils where they appear to branch or merge. These images are from a sample with 3.1M ammonium sulphate. Image b) is an enlarged area of image a). Line profiles beneath the images provide a measure of feature height (nm) along the scan line highlighted in each image.

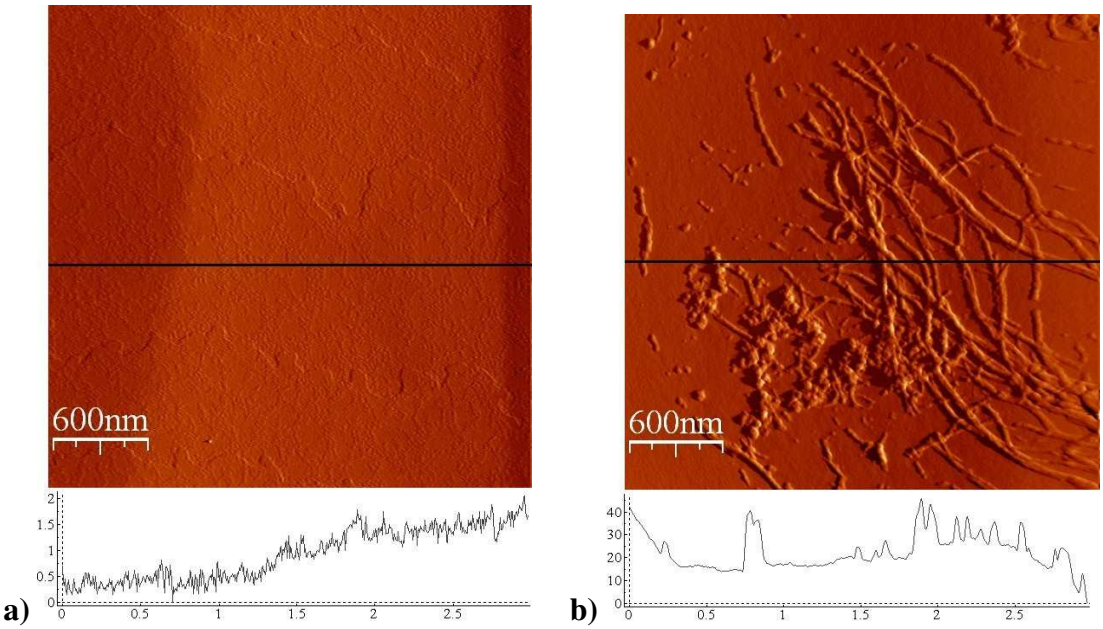


Figure 5.28: AFM images of ocr in 20mM Tris-HCl, pH 8, 3.3M ammonium sulphate at two concentrations: **a)** 0.36μM and **b)** 3.6μM. No fibrils or ocr dimers are seen in the lowest concentration and a few fibrils are being formed in a small area of the latter sample, the rest of the sample showing no fibrils. Line profiles beneath the images provide a measure of feature height (nm) along the scan line highlighted in each image.

5.5.4 Washing AFM samples

Reversibility of ocr fibril formation is one of their most interesting aspects. Salt is integral to the formation of fibrils so AFM was used to observe the result of diluting an ocr gel. A 30μM ocr, 3.3M ammonium sulphate gel was placed on an AFM slide and left for one minute. Then 100μl of buffer (20mM Tris-HCl, pH 8) was added and the excess fluid gently drained off. The sample was left to dry overnight, as usual. Figure 5.29 shows that the fibrils have completely disappeared, leaving only protein globules behind. These images are compared to the clean AFM slide in the same figure, which has an average roughness of 0.7nm, to show that the sample was not simply washed off the slide. The remaining globules are around 25nm in diameter and 2.7nm in height.

5.5.5 Effect of pH

Section 5.4 showed that below pH 5.7 ocr precipitates and therefore cannot form the fibrils observed at pH 8. AFM was used to observe this transition range of pH values. Three samples of ocr in 3.3M ammonium sulphate were prepared at pH 5.4, 5.6 and 5.8. No fibrils were observed in pH 5.4 or 5.6, only aggregates (figure 5.30). At pH 5.8 fibrils were observed. They do not appear perfectly formed and do not span the sample. Instead they grow from an area of aggregates, in a similar manner to the 3.1M sample at pH 8 (figure 5.22). More experiments are required to study higher pH values, to show the fibrils developing into a distributed network. However, the transition between precipitation and fibril formation due to protein charge, and pH is clearly demonstrated.

5.5.6 Summary

AFM results clearly demonstrate the solution conditions required for ocr gelation and fibril formation. While fibril-like objects may form above 3.1M ammonium sulphate, pH 5.7 and 3.6μM ocr concentrations they are not fully developed and do not form a space spanning network. A high enough salt concentration is required for the proteins to aggregate, while the proteins themselves must have a large enough charge to resist precipitation. Fibrils in the gel have heights that range between 2.6nm and 10.4nm, the thickness and length of the ocr dimer. Optimal conditions for a gel have been shown to occur for 36μM ocr at pH 8 and with 3.2 - 3.3M ammonium sulphate. It is within

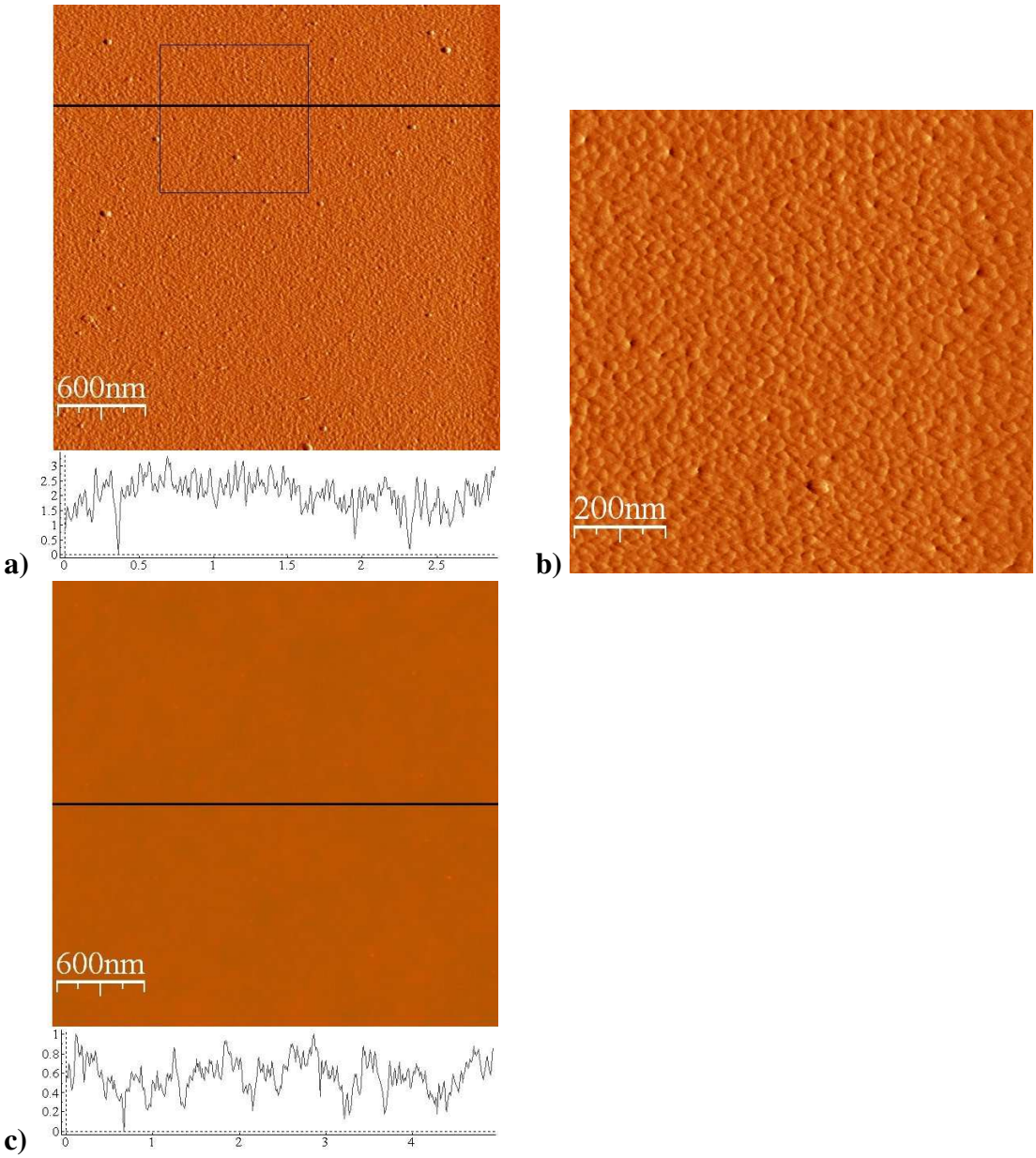


Figure 5.29: A 3.3M ammonium sulphate ocr gel was prepared on an AFM slide then washed with 100μl of 20mM Tris-HCL, pH 8. These images show that all fibrils have been dissolved (due to the removal of salt) and only globules of protein remain. Image b) shows an enlarged portion of image a), as indicated by the square. A clean AFM slide is shown in c) for comparison. Line profiles beneath the images provide a measure of feature height (nm) along the scan line highlighted in each image.

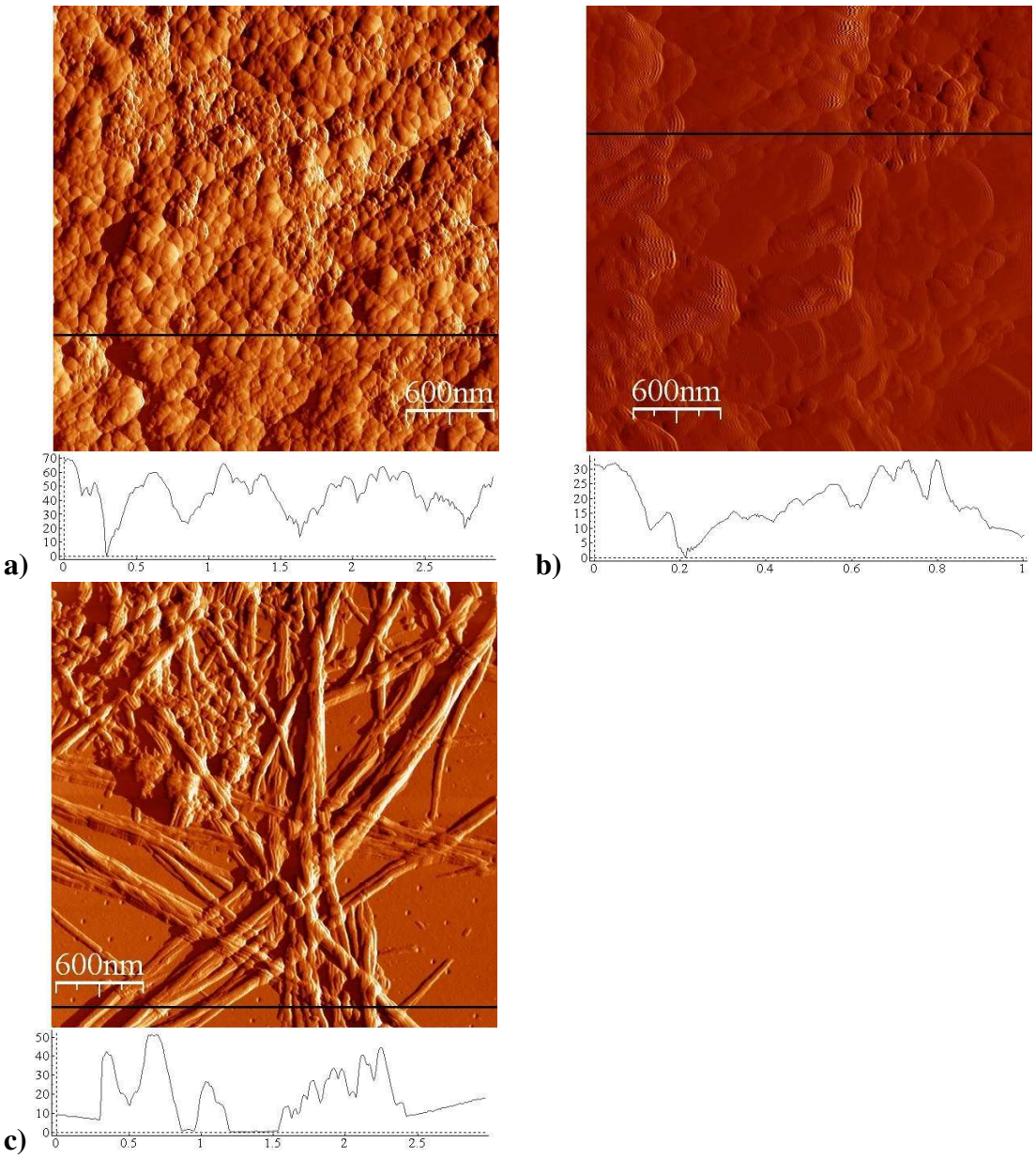


Figure 5.30: AFM images of ocr in 3.3M ammonium sulphate at three pH values around the gel transition: a) pH 5.4, b) pH 5.6 and c) pH 5.8. Fibrils begin to be formed at pH 5.8 when the charge on ocr becomes negative enough to resist precipitation. Line profiles beneath the images provide a measure of feature height (nm) along the scan line highlighted in each image.

these conditions that a model for the fibrils will be discussed in chapter 7.

Chapter 6

Secondary experiments

The three experiments described in this chapter provided useful information about the ocr gel but are not as informative as those described in chapters 4 and 5. Observing the natural diffusion of beads in ocr solutions highlighted the formation and disintegration of structure as salt concentration increased. It also raised the possibility of expanding the method to perform microrheology on the gel. Fluorescence anisotropy probed the rotational freedom of the ocr dimers but revealed only a slight increase in anisotropy with salt concentration. Fluorescence microscopy was attempted with the aim of directly observing the gel structure but the fluorescent dye used prohibited the formation of fibrils.

6.1 Particle tracking

6.1.1 Method

Surfactant-free, white, sulphate latex beads from Interfacial Dynamics Corporation UK were used for the particle tracking. They were added to the ocr sample before the dialysis so were uniformly separated. The particles were 4μm in diameter and comprised 0.4% of the final sample volume. An Olympus BX50 microscope was used to observe the beads, with a Plan 40$\times$, 0.65 Ph 2 objective. A digital camera recorded the image and the particle tracking was performed with Image-Pro software, which automatically detects particles and records their position over time. 10 areas per sample were observed each for 100 seconds. Individual particles were followed until 100-200 distinct tracks were found in total. The recorded data, showing the position of the par-

ticles at each time step, was exported to a spreadsheet and the displacement of each particle relative to its initial position calculated. The mean movement of each area was also recorded to remove effects due to vibrations or sample slippage. To do this, the average of displacement of all the particles in an area at each time step was calculated (the centre of mass displacement) and the result subtracted from each particle displacement. Finally the maximum displacement for each particle was calculated and used to build up a histogram of the number of particles with a given displacement.

6.1.2 Results

Figure 6.1 shows an image of the latex beads in a 2.0M ammonium sulphate, 100μM ocr, 20mM Tris-HCl, pH 8 solution. The particles were tracked over 100s and their position at each time recorded. The displacement of the beads in this area, relative to their initial position, is shown in figure 6.2a) and the average displacement of the beads is shown in figure 6.2b). The maximum bead displacement is around 3μm, while the common drift of the sample is an order of magnitude less. The equivalent graphs for a 3.2M ammonium sulphate, 100μM ocr, 20mM Tris-HCl, pH 8 solution are shown in figure 6.3. Here both the individual displacements and centre of mass displacement are an order of magnitude smaller than in the 2.0M ammonium sulphate sample. The movement of the particles is restricted by the formation of a gel.

Figure 6.4 shows the distribution of displacements of latex beads in five samples of 100μM ocr. For 0.0M ammonium sulphate the distribution is centered on 2μm, which corresponds to the root mean square displacement of particles due to diffusion. As the salt concentration increases the distribution becomes narrower and the rms displacement decreases to 1.6μm for 2.0M then 0.6μm for 2.8M (see inset). At 3.2M and 3.3M the distribution is tightly centered at 0.2μm. This is the pixel limit of resolution and the particles are fixed in solution. At 3.8M the distribution relaxes and the rms displacement increases back to 0.6μm. A gel transition has been detected at 3.2M once again and at high salt concentration the solution becomes less viscous. The graph of the rms displacements bears a striking resemblance to that of the shear rates for similar ocr/salt solutions as measured by rheology (figure 5.9).

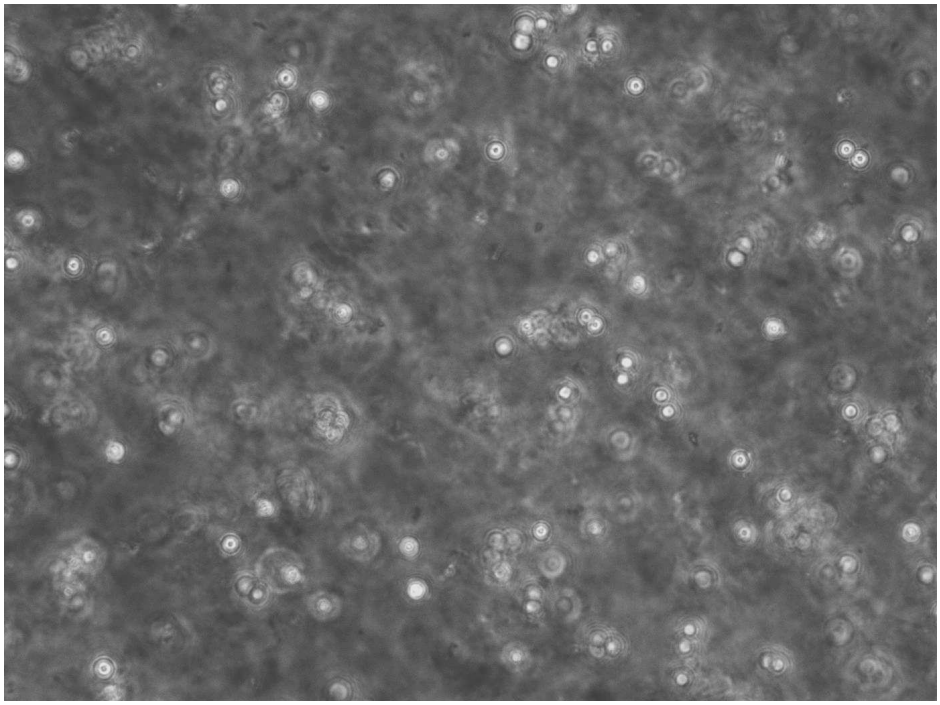


Figure 6.1: Image through the microscope of latex beads in a 2.0M ammonium sulphate, 100μM ocr, 20mM Tris-HCl, pH 8 solution. A plan 40$\times$, 0.65 Ph 2 objective was used. The scale is 160μm along the y-axis and 214μm along the x-axis.

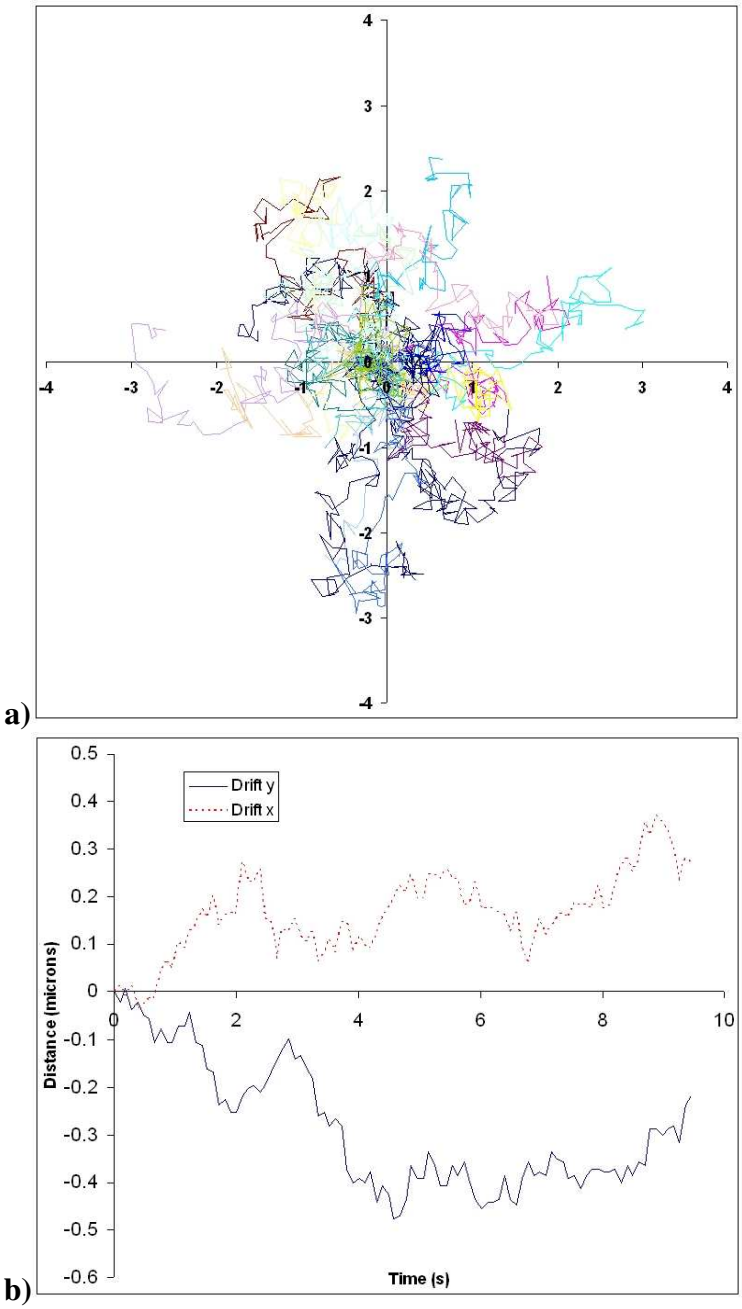
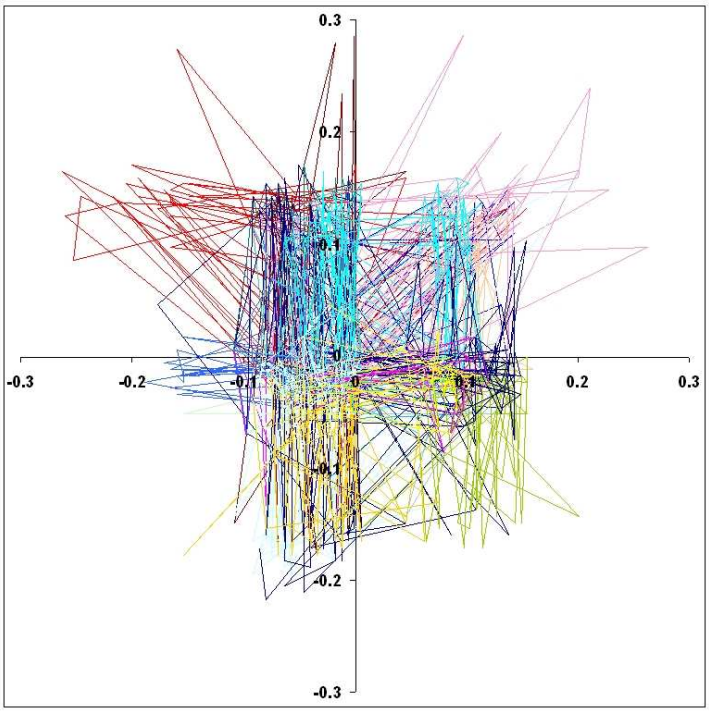
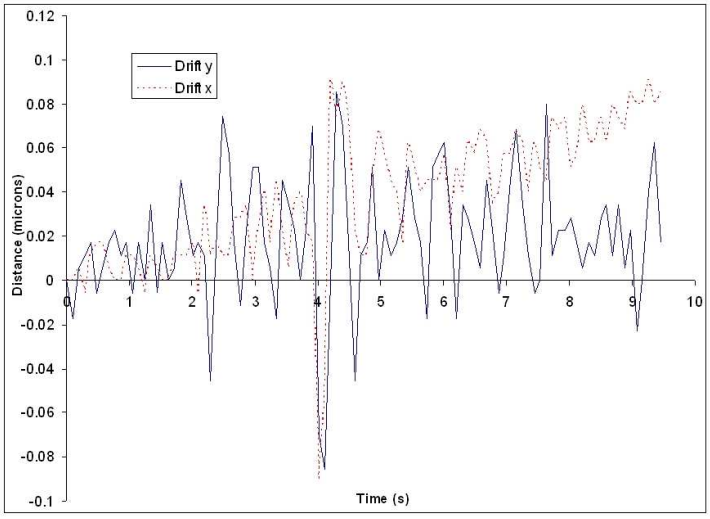


Figure 6.2: **a)** The displacement of individual latex beads from their initial position in a 2.0M ammonium sulphate, 100μM ocr, 20mM Tris-HCl, pH 8 solution for one area of observation. The axes are the x and y spatial coordinates and the scale is in μm. **b)** The average displacement of all the particles in the area, describing the drift of the solution under the microscope.



a)

(Note that the scale of the axes are an order of magnitude smaller than in figure 6.2a)



b)

Figure 6.3: **a)** The displacement of individual latex beads from their initial position in a 3.2M ammonium sulphate, 100μM ocr, 20mM Tris-HCl, pH 8 solution for one area of observation. The axes are the x and y spatial coordinates and the scale is in μm. The particles move over distances an order of magnitude smaller than in 2.0M ammonium sulphate. **b)** The average displacement of all the particles in the area, describing the drift of the solution under the microscope.

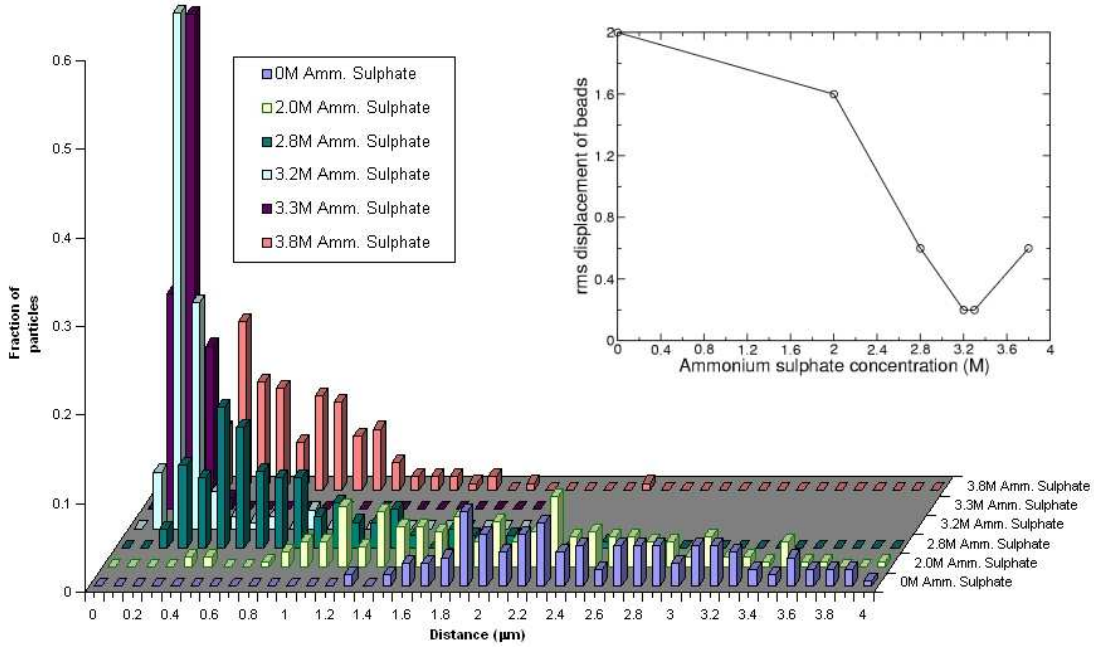


Figure 6.4: Particle tracking in six solutions of 100μM ocr and 20mM Tris-HCl (pH 8) with varying concentrations of ammonium sulphate containing 0.4% latex beads. The distribution of distances moved indicates an increasing viscosity toward the gel point at 3.3M ammonium sulphate then a relaxation as the saturation point approaches. The inset is a summary showing the rms displacement at each salt concentration.

6.2 Fluorescence anisotropy

6.2.1 Method

Fluorescence anisotropy is an important technique for studying the rotational motions of molecules in liquids. It is based on the fact that fluorophore transition moments have a defined orientation relative to the molecular axis. Fluorophores will preferentially absorb photons whose electric field vector is parallel to this moment, with a transition probability proportional to the component of electric field along the fluorophore moment. During the excited state lifetime, the fluorophore may rotate. Recording the intensity of light emitted in polarizations parallel and perpendicular to that of the exciting light provides a measure of this rotational motion: anisotropy. Anisotropy is defined as the difference between these two emission intensities. To ensure the anisotropy is independent of the total intensity of the sample it is normalized by this quantity, resulting in the dimensionless value shown in equation 6.1 [135]. For completely polarized light $I_{perpendicular} = 0$ and so $r = 1$; for unpolarized light $I_{perpendicular} = I_{parallel}$ and so $r = 0$.

$$r = \frac{I_{parallel} - I_{perpendicular}}{I_{parallel} + 2.I_{perpendicular}} \tag{6.1}$$

Steady state fluorescence spectroscopy measurements were performed on an Edinburgh Instruments FS900 fluorimeter. Ocr solutions were made with a range of ammonium sulphate concentrations. Each sample was excited with vertically polarized light (V) at 295nm (tryptophan fluorescence) and its emission intensity measured at 350nm with a vertical (VV) or horizontal (VH) polarizer in front of the detector at a bandwidth of 3.5nm. The spectra were measured with excitation and emission path lengths of 2mm and averaged over 5 scans. A buffer baseline was also measured and, although very small, was subtracted from the ocr fluorescence spectrum. Anisotropy was calculated using equation 6.1, which can be written as

$$r = \frac{VV - G.VH}{VV + 2G.VH}$$

where VV and VH are the fluorescence intensities of the vertically and horizontally polarized emission radiation. To ensure the anisotropy is independent of the total intensity of the sample ($VV + 2G.VH$) it is normalized by this quantity, resulting in a dimensionless value for r. The G factor is an artifact of the equipment. To account

for any discrepancies between the parallel and perpendicular detectors the sample is excited with horizontally polarized light and the ratio of vertically and horizontally polarized emission intensities measured [135]. The G factor is defined as this ratio:

$$G = \frac{HV}{HH}$$

6.2.2 Results

Samples were made with 30μM ocr and a range of ammonium sulphate concentrations (0.0M, 1.2M, 2.4M, 3.3M) in 20mM Tris-HCl, pH 8. Each was excited with horizontally then vertically polarized light and their fluorescence intensities measured (figure 6.5). The G-factor and anisotropy at each wavelength was calculated using these fluorescence intensities (figure 6.6). The G-factor at the emission wavelength of tryptophan (350nm) was 1 and so the anisotropies around this wavelength were comparable (see insets). An average anisotropy over the range 340-360nm was taken for each sample and plotted in figure 6.7.

As can be seen there is no major change in anisotropy across the range of salt concentrations. This indicates that the protein is not denaturing in high salt concentrations as the anisotropy would decrease to zero in that situation. The slight increase in anisotropy as the salt saturation increases probably reflects a decrease in freedom due to either the sample gelling or viscosity. The latter effect can be demonstrated by the inset in figure 6.7. Here the anisotropy, r, of the ocr solution has been related to its viscosity, ν, through the Perrin equation $\nu = \frac{\tau kT/V}{\frac{r_\circ}{r}-1}$ [135, 136], where τ is the lifetime of the excited state of the fluorophore, $r_\circ$ is the anisotropy in the absence of rotational diffusion and V is the volume of ocr. The increasing viscosity of the ocr solution, calculated from the anisotropy, matches the trend of a pure ammonium sulphate solution (data from the *"CRC Handbook of Chemistry and Physics"*).

Ocr has only one tryptophan per subunit (W94), which is on the surface. Gelation has no significant effect on the anisotropy, which shows that the tryptophan remains locally mobile on the fluorescence timescale (10^{-9}s). It does not see slow motions in the gel, only local solvent viscosity.

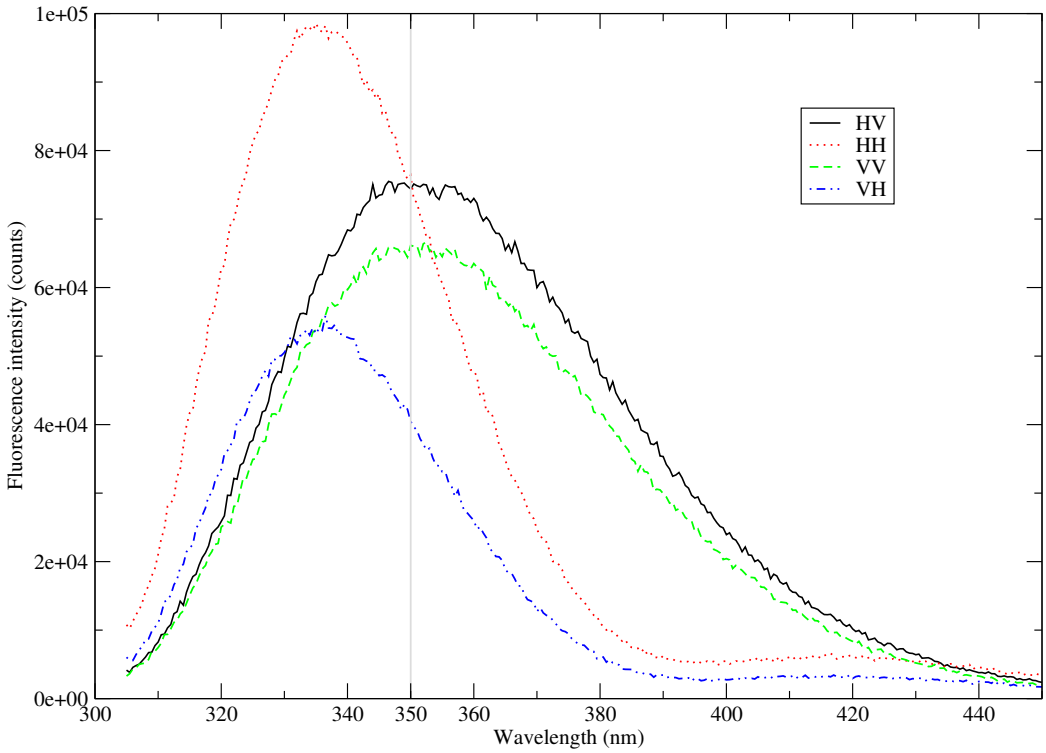


Figure 6.5: The fluorescence intensities of a 30μM ocr solution in 2.4M ammonium sulphate, 20mM Tris-HCl, pH 8. HV and HH represent the sample excited by horizontally polarized light then measured through a vertical and a horizontal polarizer respectively. Similarly VV and VH represent the sample excited by vertically polarized light. The former pair are used to calculate the G-factor and the latter pair the anisotropy. The emission wavelength of tryptophan (350nm) is highlighted.

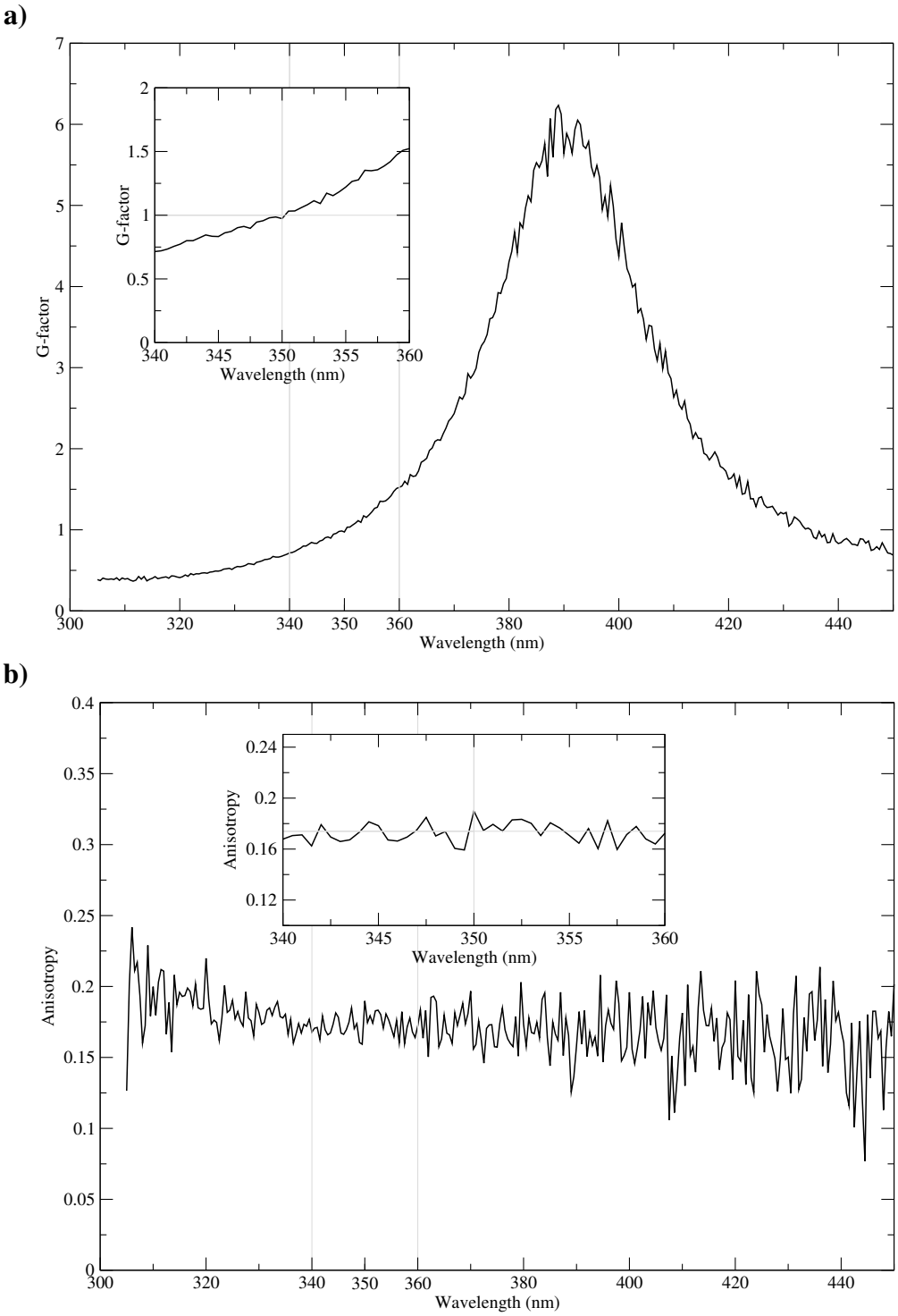


Figure 6.6: The G-factor **(a)** and anisotropy, r, **(b)** at each of the measured wavelengths for a 30μM ocr solution in 2.4M ammonium sulphate, 20mM Tris-HCl, pH 8. The inset in each graph is an enlargement of the data over the emission wavelength of tryptophan (340-360nm). The quantities are calculated from the fluorescence intensities (figure 6.5) as follows: $G = \frac{HV}{HH}$ and $r = \frac{VV - G.VH}{VV + 2G.VH}$

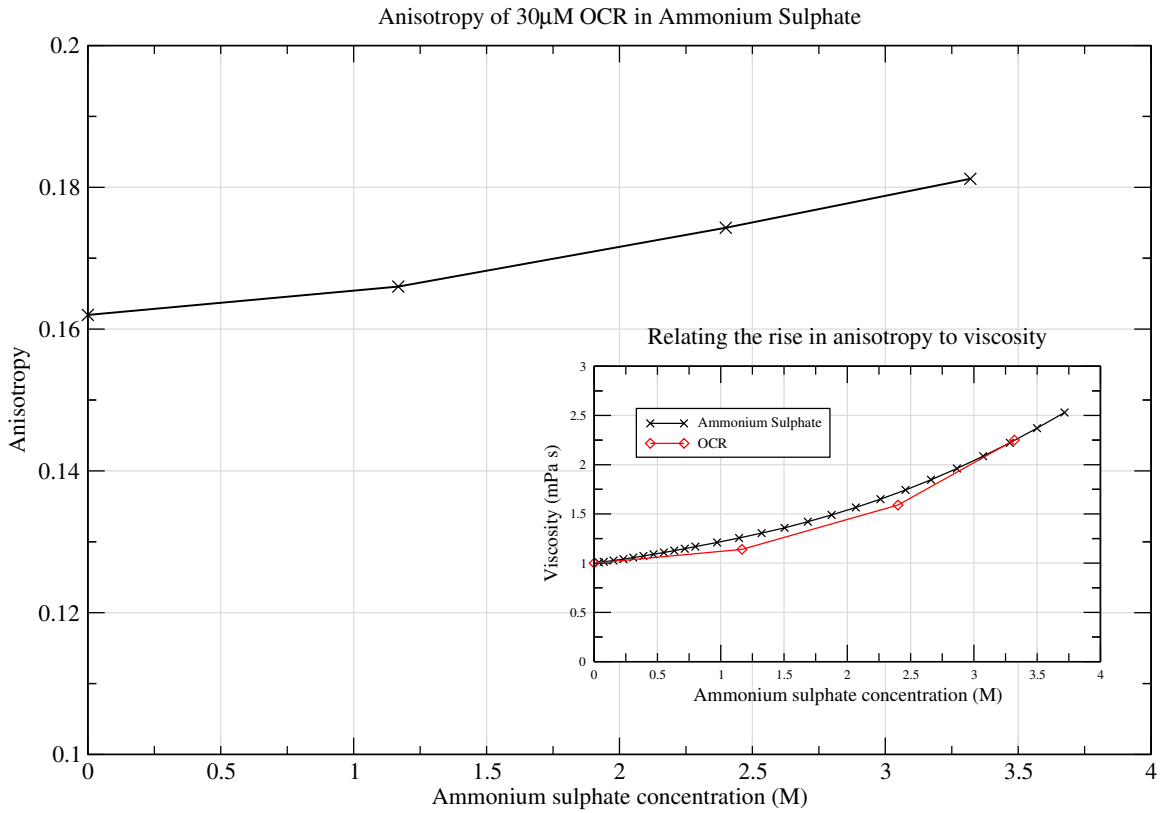


Figure 6.7: Fluorescence anisotropy of a 30μM ocr solution with a range of ammonium sulphate concentrations. There is no obvious trend save the gradual increase with salt saturation. This can be related to the increase in viscosity due to high concentrations of salt (inset). Specifically the gel point has no effect on anisotropy.

6.3 Fluorescence microscopy

6.3.1 Method

Confocal microscopy was used to examine the underlying structure of the gel. A mutant form of ocr with a surface aspartic acid replaced by a cysteine amino acid was used [55, 62]. The ocr mutants D25C and D62C were used for labelling, referring to the replacement of the 25th or 62nd aspartic acid (D) on the amino acid chain with a cysteine (C) residue. This cysteine was labeled with a 5-IAF[1] fluorescent dye from Molecular Probes UK. 5-IAF is highly water soluble, it has a charge of -3 and its structure is shown in figure 6.8. It reacts with the thiol group (sulphur-hydrogen) on the cysteine amino acid to form a thio-ether bond between molecules. To label ocr the sample was incubated overnight at 4°C in the dark with a 50-fold molar excess of dye. β-Mercaptoethanol was added to 100mM to stop the reaction and unbound dye was removed by dialysis against 20mM Tris-HCl, pH 8. UV spectra from before and after labelling were compared to calculate the percentage of protein labelled with dye. The concentration of bound dye was calculated as 782μM and the concentration of ocr was 433μM. If there are two dyes per dimer, this means that 90% of ocr proteins were labelled. Aliquots of the labeled sample were dialyzed against a range of ammonium sulphate solutions to give four 50μl solutions of 36μM ocr in 0M, 2.4M, 3.3M and 3.8M ammonium sulphate. These were then observed with a Bio-Rad confocal microscope using an argon ion laser to excite 5-IAF at 488nm and emission measured at wavelengths over 500nm (5-IAF emits maximally at 515nm). A PF $\times$40 oil interface lens was used and images recorded at 500 lines per second over an area of 308$\times$308μm.

6.3.2 Results

The labelled ocr solutions were prepared with four different concentrations of ammonium sulphate: 0.0M, 2.4M, 3.3M and 3.8M. For each one, images were taken for a time series in one plane with an interval of 3s, and a z-stack over 100μm in 1μm steps. Figure 6.9 shows an example image taken for each sample. Part a) shows that ocr in buffer only does not form any kind of structure. More importantly it shows that the dye does not cause ocr to coagulate.

[1] 5-iodoacetamidofluorescein

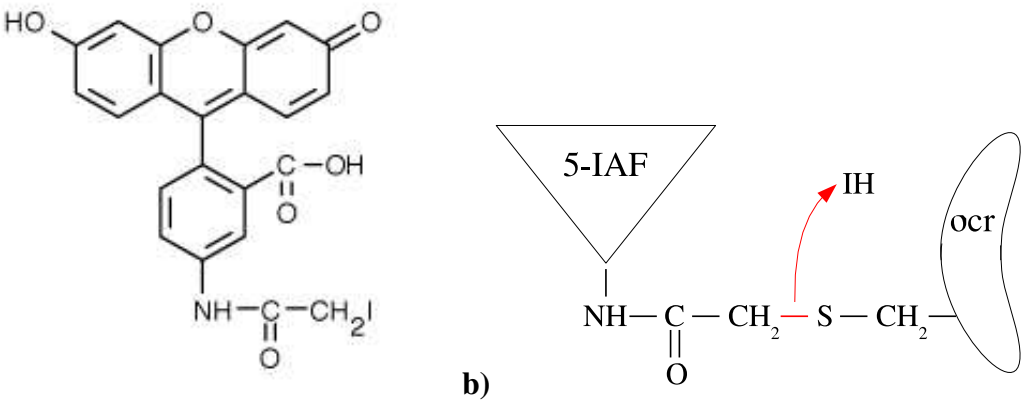


Figure 6.8: **a)** Structure of the fluorescent label 5-iodoacetamidofluorescein (5-IAF). It has a charge of -3 and a molecular weight of 515.26. Its molecular formula is $C_{22}H_{14}INO_6$. **b)** The label binds to the cysteine residue on ocr by displacing the iodine on the dye and replacing it by forming a bond with the sulphur on cysteine.

At a fixed time and in a fixed plane the other three images look similar in style. The ocr molecules have aggregated into distinct, little groups. They are not globules but have an extended structure of their own. In part b) they appear more uniformly distributed and the z-stack shows that they do not percolate. In contrast the z-stacks of 3.3M and 3.8M solutions show the structure percolating across the sample. This agrees with the observation that a gel is formed above 3.2M ammonium sulphate. The time series images reveal the weakening of this gel above 3.4M as diffusion of aggregates are observed. Images a) and b) show proteins and aggregates freely moving through the sample. In image c) no movement is observed at all, while in image d) small aggregates are seen to travel across the image. These results support those from circular dichroism and particle tracking.

Although visual observations of viscosity are reproduced by confocal microscopy it is not clear that the structures observed are actually those found in gels of ocr without fluorescent labels attached. Aggregates are observed in the images on the same scale as the wavelength of light so the gel should not be transparent, in contrast to the gel obtained with unlabelled ocr. Indeed, the transmitted confocal microscope image of the 3.3M sample (figure 6.10a) implies a turbid material has formed. It bears a likeness to the image of a turbid, transient lysozyme gel in figure 6.10b). However, only small samples were available for use with the microscope and these were highly fluorescent

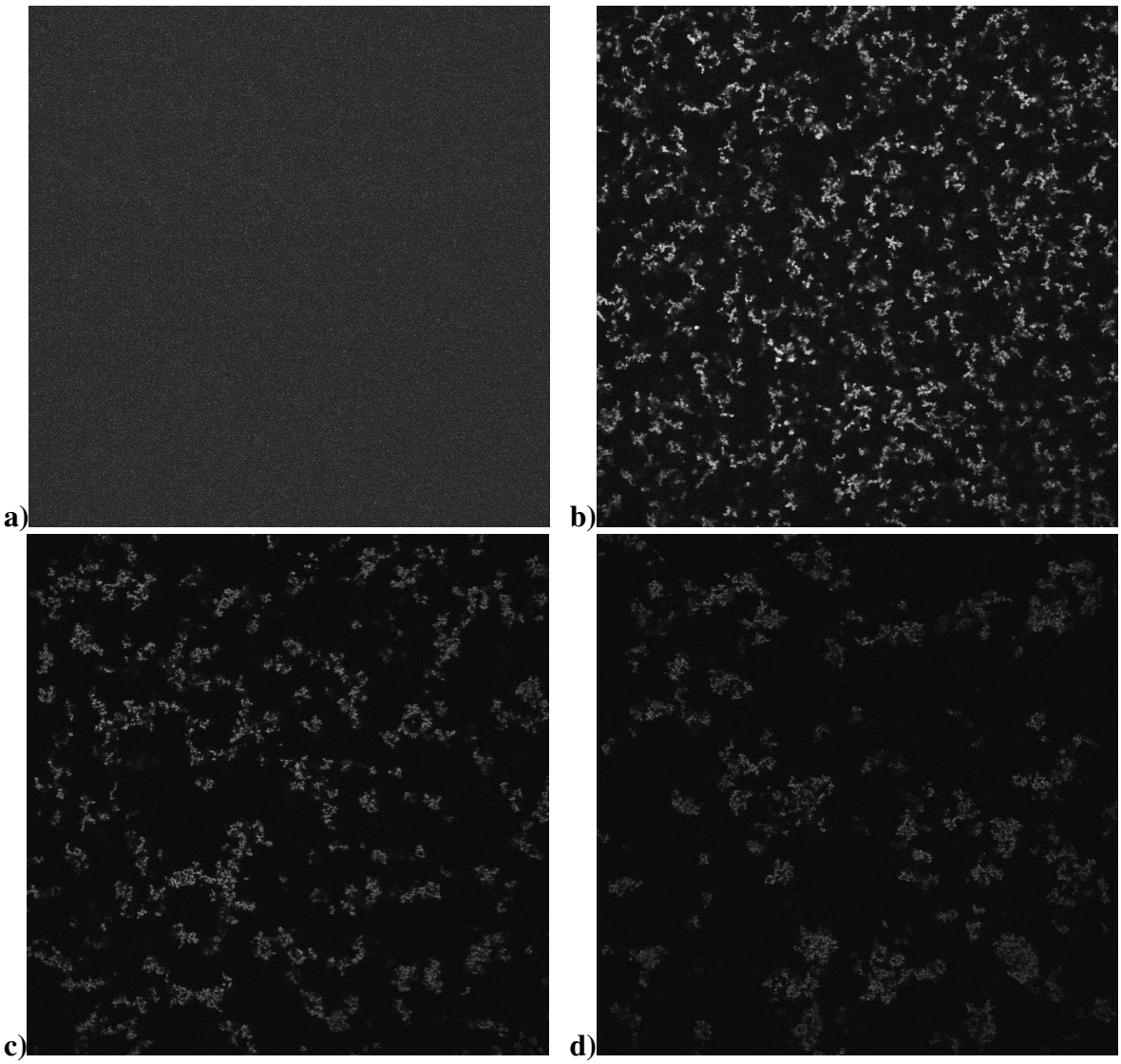


Figure 6.9: Confocal fluorescence microscopy images of labelled ocr solutions in ammonium sulphate concentrations of a) 0.0M, b) 2.4M, c) 3.3M and d) 3.8M. The size of each image is $308 \times 308 \mu$m. The sample forms a gel above 3.2M, which weakens above 3.4M ammonium sulphate.

so it was not obvious to the eye whether they were transparent or not. Furthermore the volume fraction used in the samples is sufficiently low (~1%) that turbidity may not be observable with the naked eye.

5-IAF has a charge of -3 (figure 6.8) and there are two labels per ocr dimer. Since an aspartic acid residue was removed from each monomer the net change is -4e per dimer, making ocr even more negative. As described in chapter 7 our model for gelation relies strongly on each negatively charged amino acid on the surface of ocr. Altering the charge could significantly affect the process. The size of the dye may also be a prohibitive factor, affecting the ability of ammonium sulphate to bridge between proteins. It is still possible that a gel would form but it may be simply due to aggregates forming a space-spanning structure, which would be turbid, rather than polymers forming a network, which would lead to a transparent sample. If only the introduced cysteine residues are labelled, this may indicate that these sites play an important role in the fibril formation process. However, two different sites were used so it cannot be that specific. It may be that the dye is binding to other residues and some proteins are gaining more than two lables, while others a left un-labelled. A simple experiment to check this would be to add the dye to a native solution of ocr, with no cysteines present, and see if any of the proteins are labelled.

6.3.3 AFM

Atomic force microscopy (chapter 4.3) of labelled ocr was performed in three salt concentrations, spanning the gel point: 3.0M, 3.2M and 3.4M. Fibrils were not observed in any of the samples. It appears that addition of the 5-IAF label prohibits both the production of fibrils and the formation of a gel.

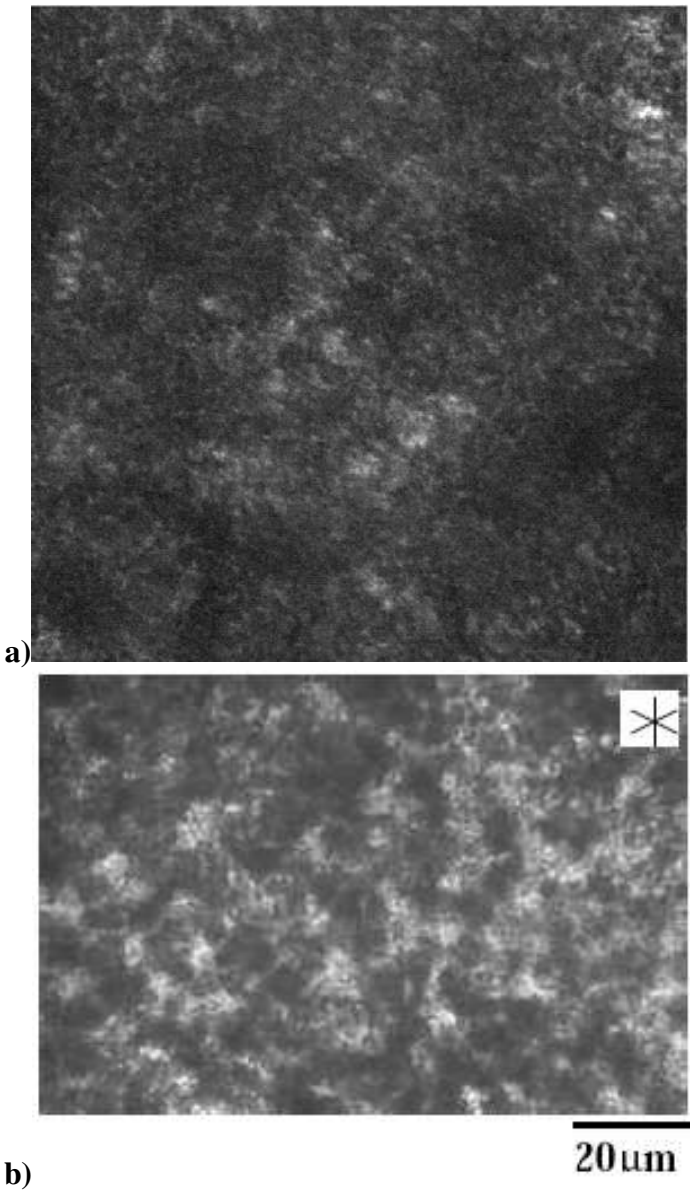


Figure 6.10: **a)** Confocal microscopy image of labelled ocr solution in 3.3M ammonium sulphate. A turbid solution has formed with the addition of salt. The size of the image is $120 \times 120 \mu$m. **b)** Optical microscope image of a lysozyme gel (volume fraction $\phi = 0.1$) formed in 0.9M NaCl at 22°C and pH 4.5. Image reproduced from [73].

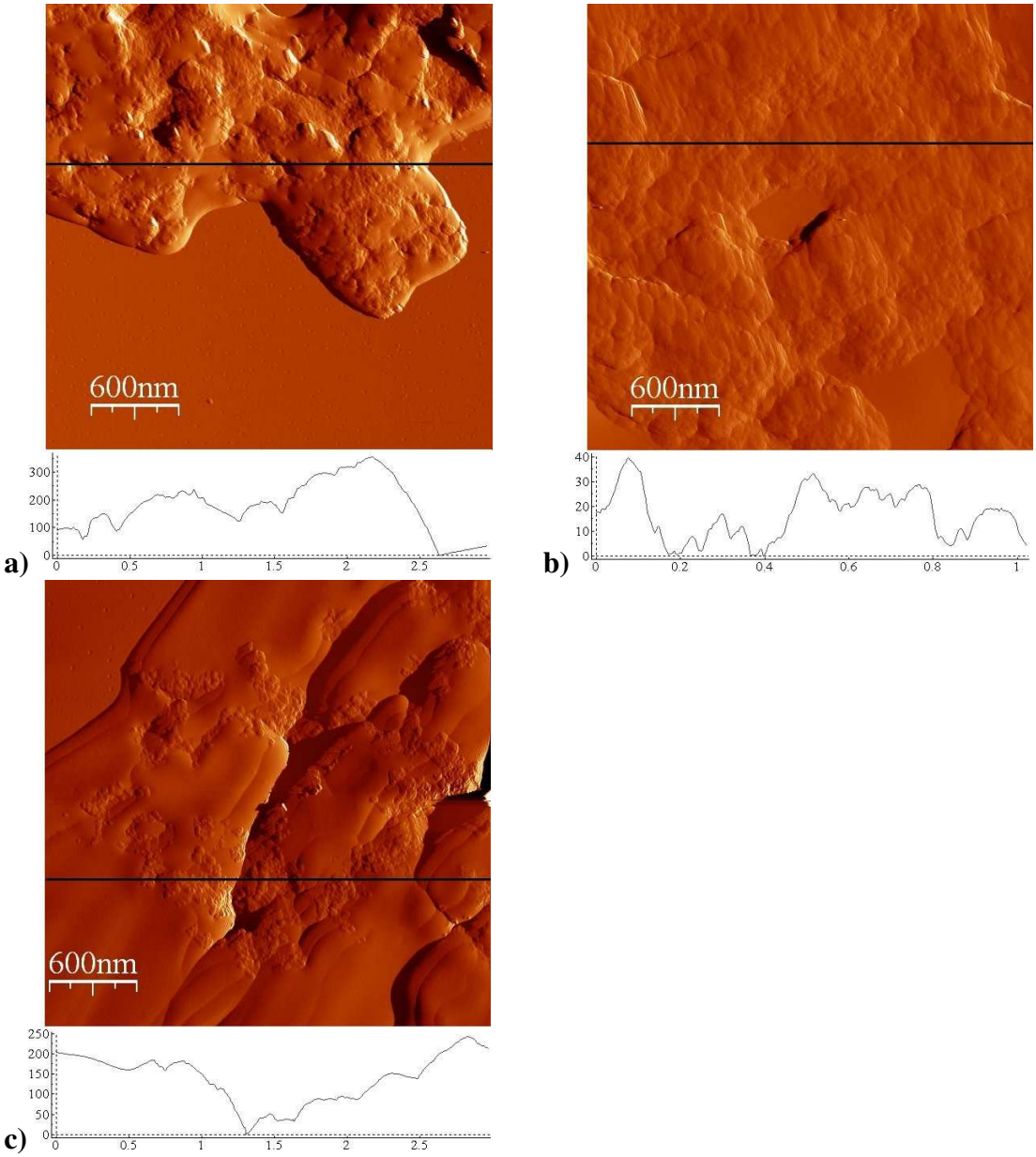


Figure 6.11: AFM images of 36μM ocr labelled with 5-IAF in solutions of 20mM Tris-HCl, pH 8, and **a)** 3.0M **b)** 3.2M **c)** 3.4M ammonium sulphate. None of the samples (at, above or below the gel point) show evidence of fibrils. The label causes the proteins to aggregate regardless of the salt concentration, and before fibrils may be formed.

Chapter 7

A model for ocr fibrillation

7.1 Preliminary observations

Ocr forms a gel upon the addition of a specific salt, without any heating. Thus, rather than simply screening the protein charges, ammonium sulphate must cause aggregation by binding to the proteins. However, it does not cause the protein to denature, alter its secondary structure or precipitate at physiological pH. The gel forms rapidly but only above a very high ammonium sulphate concentration (3.2M), and it then appears to deteriorate as ammonium sulphate tends towards saturation. Although the salt must initially act to stabilise the protein, rather than precipitating it, the increased concentrations may eventually cause the protein structure to unfold (as shown by the analysis of CD spectra, figure 5.13).

Since the gel is transparent the underlying structure must have at least one dimension smaller than optical wavelengths, but it must also be spatially extended in one or more dimensions to form at low protein concentrations. AFM revealed that ocr was forming long, stiff fibrils, in agreement with these conditions. AFM also provided a number of observations that has to be reproduced by any model: these included branching, bending and a fibril thickness lying predominantly between 2.6nm and 10.4nm.

7.2 An entropy driven model

At such high salt concentrations the ratio of cations to negative ocr charges is in the region of 1000:1 (table 5.2). The formation of a gel cannot simply be ascribed to charge

screening though, as demonstrated by ocr's disregard for all salts other than ammonium sulphate. There are numerous systems in nature where the screening ability of a salt is of less importance than its chemical properties. For example, DNA condensation behaviour is strongly dependent on the specific binding of ions as well as their valency. Here the divalent cations Mn^{2+} and Cd^{2+} condense DNA but Ca^{2+} and Mg^{2+} do not. The former ions adsorb in grooves of DNA, while the latter bind to the backbone of DNA [137]. However, the aggregation of ocr is not due to solely the ammonium nor the sulphate ion either. It appears to be specifically dependent on "molecular" ammonium sulphate alone.

At low concentrations of ammonium sulphate the ions dissociate in solution. Towards 1M molecular dynamics simulations [138] indicate that ammonium sulphate has a large cation-anion correlation so the ions strongly cluster into aggregates that exclude water. As the concentration increases still further the entropy decreases and to minimise the free energy the majority of ions reassociate. The ammonium sulphate aggregates may form ionic bridges between neighboring ocr proteins via the carboxyl groups on the negative aspartic and glutamic amino acids (as shown in figure 7.1). This mechanism is capable of random aggregation as well as stereo-specific binding. Indeed, amorphous ocr aggregates have been observed coexisting with poor quality fibrils at 3.1M and 3.5M ammonium sulphate (see AFM images in figure 5.22). The ocr clusters have a size of the order 300nm, lower than visible spectrum wavelengths, which is consistent with transparent solutions. An additional, as yet unknown, mechanism occurs at 3.2M that causes a specific set of NH—OC hydrogen bonds to form between the salt and the proteins. This gives rise to a well-defined orientation between ocr dimers and the consequent growth of helical fibrils that form a space-spanning, viscoelastic network. Since the structure of ocr is unperturbed during this mechanism for fibril formation, individual dimers can be regained by diluting the solution and reversing the process. As the bulk concentration of ammonium sulphate decreases, the entropy rises and the ions disassociate leaving a dispersion of free ocr dimers. This was demonstrated with AFM (figure 5.29) where diluting the gel removed the salt bridges between ocr proteins and consequently the fibrils broke apart.

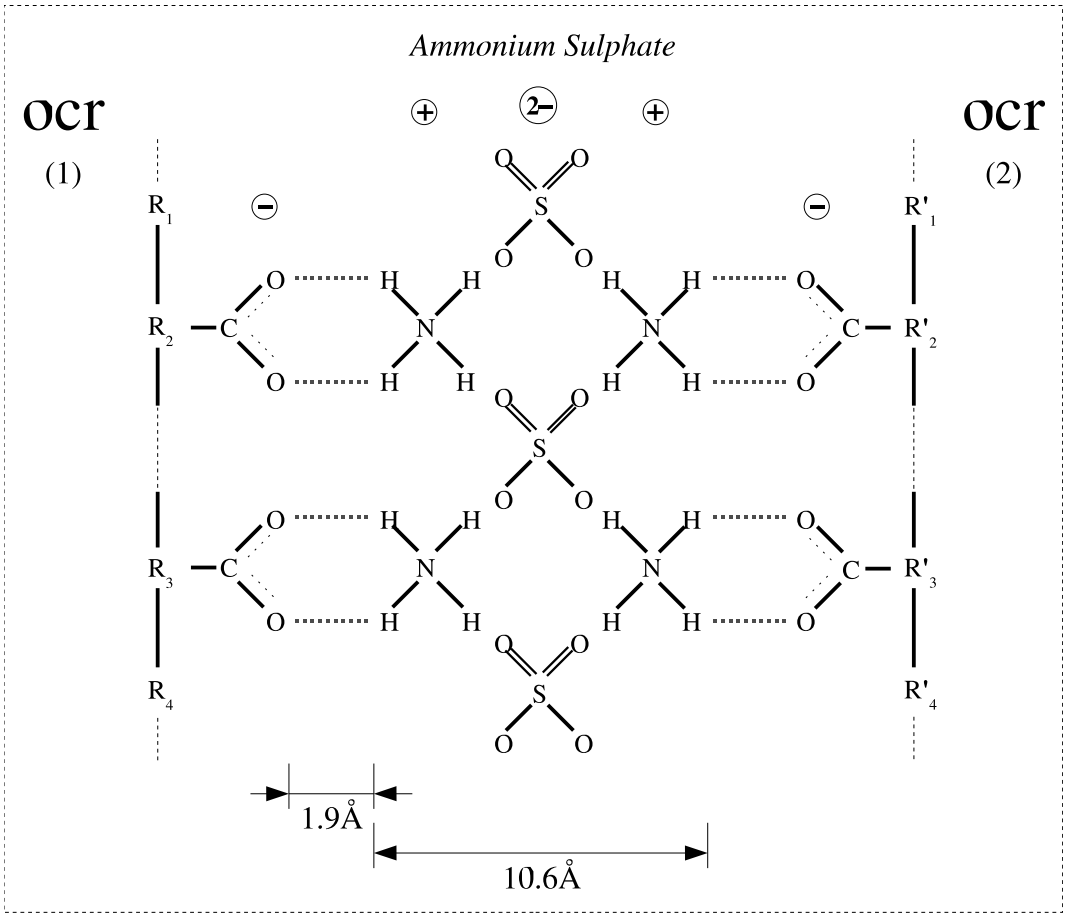


Figure 7.1: Schematic of the proposed model for the specific aggregation of ocr dimers. Ammonium sulphate molecules attach to carboxyl groups on aspartic and glutamic amino acids (R_2, R'_2 and R_3, R'_3) and form bridges between neighboring ocr dimers via NH—OC hydrogen bonds. Each R represents an amino acid in the protein's sequence.

7.3 Protein precipitation and pH

Understanding why ocr does not precipitate is an important step towards a model of gelation. As a comparison we looked at the salt-induced aggregation of lysozyme studied by Poon *et.al.* [139, 140]. By normalizing the salt concentration with the square of the protein charge they discovered a universal crystallization boundary, independent of pH. Treating the crystallization boundary as an effective salting out curve we can compare the precipitation point of lysozyme to that of ocr.

7.3.1 Lysozyme aggregation and crystallization

7.3.1.1 Lysozyme phases

The phase diagram for lysozyme in solutions of sodium chloride at pH 4.5 was measured by Sedgwick *et.al* [140] and is reproduced in figure 7.2. At low volume fractions (Φ) four non-equilibrium regimes are observed. As the salt concentration is increased the homogeneous samples begin nucleating above a crystallization boundary. When the salt concentration is high enough a non-equilibrium boundary is passed and all the samples become turbid. Just above this boundary spinodal decomposition causes a liquid-liquid phase separation and a bicontinuous fluid forms (figure 7.2a). At lower volume fractions and increasing salt concentrations amorphous aggregates are observed (figure 7.2b). These range from non-coalescing spherical particles a few micrometers in diameter to larger, irregular aggregates. At the highest salt concentrations white, transient gels form. They contained inhomogeneities on the order of 10μm (corresponding to around 10^3 particles), which can be seen in an optical microscope (figure 7.2c). The gels were not permanent, however, and rapidly sedimented after 15-30 minutes.

To investigate the relation between charge and salt concentration for lysozyme, Poon *et al.* [139] measured the phase diagram at pH 4.5, 5.9 and 7.8 (figure 7.3a). At these pH values the equivalent charge on lysozyme is +8e, 9.4e and 11e. At each pH the solutions are homogeneous until enough salt is added that van der Waals attractions become the dominant force and the proteins crystallize. As the pH increases, though, less salt is required for crystallization; aggregation begins at a lower c_s. If the data is scaled by the charge on lysozyme at the respective pH a boundary appears, separating crystal phases from single phases for all pH values. Figure 7.3b) shows the phase

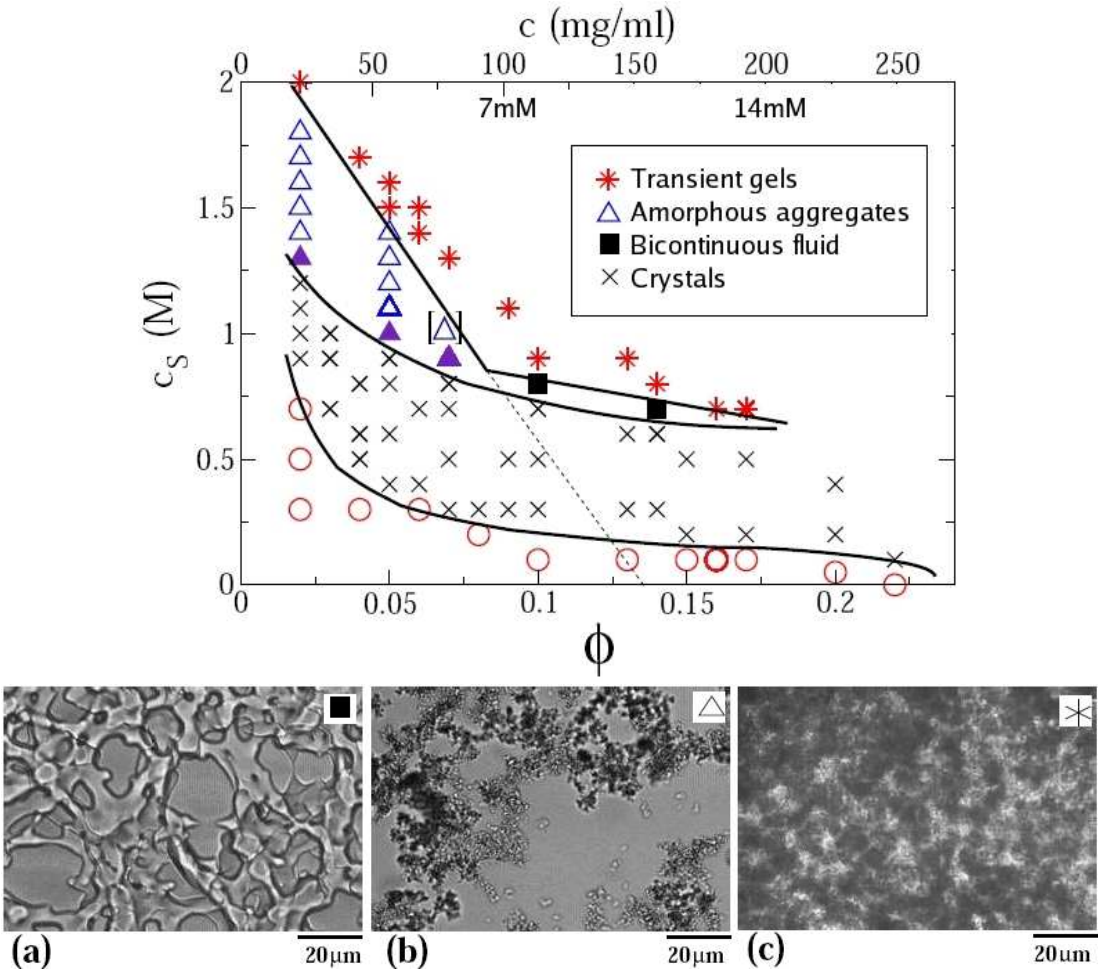


Figure 7.2: Experimental phase diagram of lysozyme in NaCl at 22°C and pH 4.5. The salt concentration, c_s, is plotted against lysozyme volume fraction, ϕ, and concentration,c. Phases are: ○ homogeneous solution, × crystals, ■ bicontinuous fluid, △ ▲ amorphous aggregates, ∗ transient gels. Optical microscope images of three phases are shown underneath: (a) bicontinuous fluid (ϕ=0.14, c_s=0.7M), (b) aggregates (ϕ=0.02, c_s=1.5M), (c) transient gel (ϕ=0.1, c_s=0.9M). For comparison, at 1mg/ml (36μM) ocr has ϕ=0.001 and it forms a gel at c_s=3.2M. 7mM ocr corresponds to c_{ocr} =190mg/ml. Reproduced from [73].

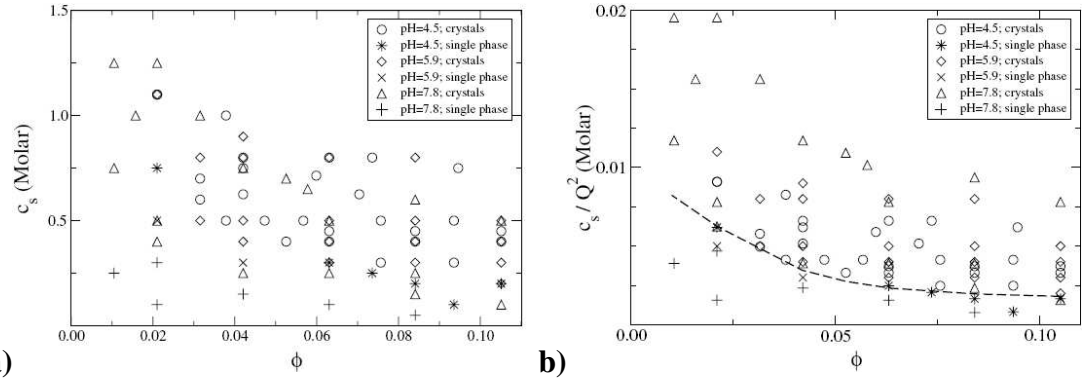


Figure 7.3: **a)** Phase diagram of lysozyme and sodium chloride solutions showing salt concentration, c_s, against volume fraction, ϕ, for pH 4.5, 5.9 and 7.8 (at 22°C). **b)** The same phase diagram scaled by Q^2. The salt concentration is normalized by the protein charge at the respective pH (c_s/Q^2). The universal crystallization boundary, below which crystals coexist with solution, is shown as a dashed line. Reproduced from [139].

diagram replotted with the salt concentration normalized by the square of the protein charge (c_s/Q^2). The "universal crystallization boundary" separating phases, is represented by the dashed line. Above the boundary crystals coexist with solution, while below only single-phase solutions are found.

The crystallization boundary and c_s/Q^2 scaling can be investigated with the second virial coefficient. The ideal gas law of a many-body system can be expanded as a power series in the density, known as a virial expansion:

$$\frac{p}{k_B T} = \rho + B_2\rho^2 + B_3\rho^3 + \ldots \tag{7.1}$$

The virial coefficients B_i are characterised by the many-body inter-particle potentials. In particular, the second virial coefficient B_2 is determined solely by the pair interaction between particles. It can be expressed as equation 7.2 for particles separated by r, interacting via the DLVO potential $\Psi(r)$ (from equation 2.10). If the hard sphere potential is factored out, leaving a potential $u(r) = \psi_V(r) + \psi_C(r)$, the reduced second virial coefficient is given by equation 7.3.

$$B_2 = 2\pi \int_0^\infty (1 - e^{-\Psi(r)/k_B T}) r^2 dr \tag{7.2}$$

$$b_2 = 1 + \frac{3}{8R^3} \int_{2R}^\infty (1 - \exp^{-u(r)/k_B T}) r^2 dr \tag{7.3}$$

In a dilute protein solution B_2 is used to look at the interactions between proteins and can be measured by static light scattering. When B_2 is positive the predominant force between proteins is repulsive so they are more likely to interact with the solution than each other. As B_2 becomes negative attractions between proteins become stronger. Crystallization at constant protein volume was shown to happen only when B_2 lies within a narrow range, for all proteins and precipitants studied by George *et al.* and Zukoski *et al.* in the (Φ,B_2) plane [141, 142]. Poon *et.al.* expanded this to show there exists an equivalent region in the (c_s, Q^2) plane, within which the conditions for protein crystallization are most favourable. They predict a common c_s/Q^2 crystallization boundary for any protein and salt.

7.3.2 Ocr and lysozyme

To look at the precipitation of ocr in terms of the universal crystallization boundary we need to compare ratios of c_s/Q^2 for ocr and lysozyme. The above discussion was concerned with the monovalent salt sodium chloride, Na^+Cl^-. For higher valency salts we need to consider the ionic strength, I, of each salt instead of just its concentration, c_s (equation 2.9). For monovalent salts the ionic strength is equivalent to the salt concentration, $I = c_s$. For ammonium sulphate, $(NH_4^+)_2SO_4^{-2}$, the ionic strength is $I = \frac{1}{2}(2.c_s.1^2 + 1.c_s.2^2) = 3c_s$.

At pH 8 the charge on lysozyme is $Q = 8$, while ocr has a charge of $Q = 56$. To precipitate ocr we would need an ammonium sulphate concentration of

$$\frac{I^{lys}}{Q_{lys}^2} = \frac{I^{ocr}}{Q_{ocr}^2} \implies c_s^{NH_4SO_4} = \left(\frac{56}{8}\right)^2 \left(\frac{1}{3}\right) = 16.3M \tag{7.4}$$

This is over four times the concentration of a saturated solution of ammonium sulphate (4M). So we should not expect ocr to precipitate under the conditions used in experiments. At low volume fractions, aggregation is observed to occur for values of y greater than 0.01, where y is defined as the ratio of solvent ionic strength to the square of protein charge:

$$y = \frac{I}{Q^2} \tag{7.5}$$

The salt concentration required to precipitate proteins can be estimated using this ratio, where $y = 0.01$ is taken as the minimum value required for aggregation. For lysozyme

Salt	Saturation (M)	Ionic strength (M)
$(NH_4)_2SO_4$	4.0	12.0
Na_2SO_4	3.2	9.6
K_2SO_4	1.4	4.2
NH_4Cl	16.3	16.3
NaCl	6.5	6.5
$MgCl_2$	3.13	9.4
$MgSO_4$	3.03	12.1

Table 7.1: Ionic strengths of various salts at their saturation concentration (at room temperature). The estimated ionic strength required to precipitate ocr is ~30M. None of the above salts have any effect on ocr, except ammonium sulphate.

$I \simeq 0.6$M, well within achievable salt concentrations, while for ocr $I \simeq 31$M, unachievably larger. The ionic strengths of saturated salt solutions are given in table 7.1. As expected, they are all significantly lower than the required concentrations to precipitate ocr, and as shown in section 5.3 none have any affect on the protein, except ammonium sulphate. Ammonium sulphate is still unable to precipitate ocr but at 3.2M ($I = 9.6$M) it forms a gel with ocr via a different interaction, which is discussed later.

If the charge on ocr is the controlling factor under high salt conditions then reducing the charge should cause it to precipitate. We can control the charge on ocr by varying the pH of the buffer. To precipitate ocr at 3.2M it needs a charge of:

$$\begin{aligned} \left(\frac{Q_{ocr}}{8}\right)^2 \left(\frac{1}{3}\right) &= 3.2M \\ Q_{ocr} &= \pm 24.8 \end{aligned}$$

The pH to which this is equivalent can be found from the titration curve in figure 2.17a) but depends on the model used. The theoretical titration curve (based on individual pK_a values) gives a pH of 4.4, while Bo Jönsson's simulations predict a pH of 5.7 for an ionic strength of 3mM and pH 4.7 for 470mM. Although these are much lower than experimental ionic strengths, the simulations predict a trend towards the theoretical titration curve. This is in agreement with the UV absorption spectra in section 5.4, which indicate that ocr precipitation begins around pH 5.6 and by pH 4 the protein has almost completely precipitated.

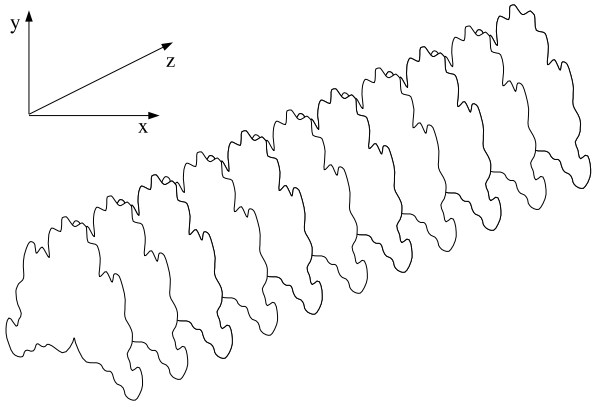


Figure 7.4: A scale diagram of ocr dimers forming a fibril by aligning side by side. An argument against this model is that it restricts the freedom of tryptophan residues on the sides of the dimer, which is not observed in tryptophan fluorescence anisotropy experiments (figure 6.7).

7.4 Fibril structures

Perhaps the simplest configuration of ocr dimers in a fibril is a side-by-side alignment, as shown in figure 7.4. This reproduces the observed cross-section dimensions of 2.6nm$\times$10.4nm and allows the fibril to bend. However, it is not clear how branching would occur in this arrangement. Another criticism of side-by-side aggregation involves the location of the two tryptophan amino acids. These are on the sides of the dimer (figure 7.5) and their freedom of movement would be affected by neighbouring dimers. However, the tryptophan fluorescence anisotropy showed little change upon gelation (figure 6.7), implying that a side-by-side aggregation may not be as likely as other possible orientations.

7.4.1 Top-to-bottom stacking

Another likely orientation of ocr dimers in a fibril is top-to-bottom stacking, along the bow axis of the protein. A scale diagram showing a plausible alignment of dimers in two dimensions is given in figure 7.6a). The ocr dimer is represented by the outline of its crystal structure (figure 2.8b), a more realistic approximation than a symmetric banana model as the top and bottom of the protein's outline do not have the same curvature. This simple model depicts fibrils as flat tapes. However, it is possible, if not probable, that the dimers align with a slight twist out of the plane of the paper.

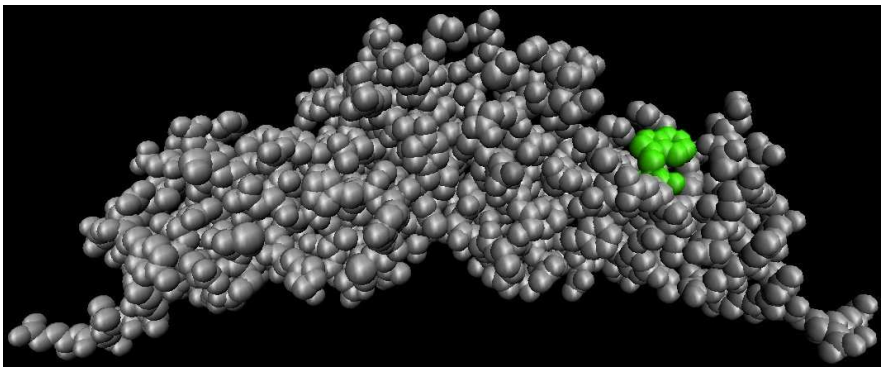


Figure 7.5: Location of a tryptophan amino acid on one side of an ocr dimer. There is a second residue on the rear side, belonging to the left-hand monomer.

The prolate ellipsoid model for the ocr dimer (figure 2.6) gives it a thickness of 2.6nm but the banana shape from x-ray crystallography of the protein allows for a closer packing. As a result the distance two dimers cover in this model fibril is 5nm, rather than 5.2nm. The cross-sectional dimensions of the proposed fibril are 2.6nm$\times$10.4nm (the thickness and length of the ocr dimer), which matches the measured fibril height distribution for a 3.2M ammonium sulphate gel (figure 5.23). Assuming the dimers will bind to each other over the largest surface area available in this orientation, an angle of around 18° is required between the proteins. This tilt between dimers means a minimum period of 4 proteins is required for the fibril to remain in a straight line (figure 7.6b), albeit with a slight ripple along the sides. It also allows a second dimer to bind in the gap between the original two proteins. No specific links between groups are proposed but it is clear from figure 7.7b) that the location of charged groups is conducive to binding in this arrangement. A new fibril may grow from this protein, with dimers continuing to bind at 18°. A predicted branch is shown in figure 7.7a). An angle of 54° is produced between the straight-line paths of the original fibril and the branch. This matches a branch angle observed in an AFM image of ocr fibrils, seen in figure 5.27. However, forming a junction in this manner involves a different binding orientation from the fibril core. Binding along the fibril takes place between the same monomers on adjacent dimers; either right with right or left with left. The hypothesised branch relies on binding between opposite monomers, eg. right with left. This is possible if it occurs between the acidic residues, arranged on the surface as a helix, as shown in figure 7.7b). Other branch angles have been observed in ocr fibrils (figure 7.8) and range in angle from around 20° - 60°. This may be due to flexibility

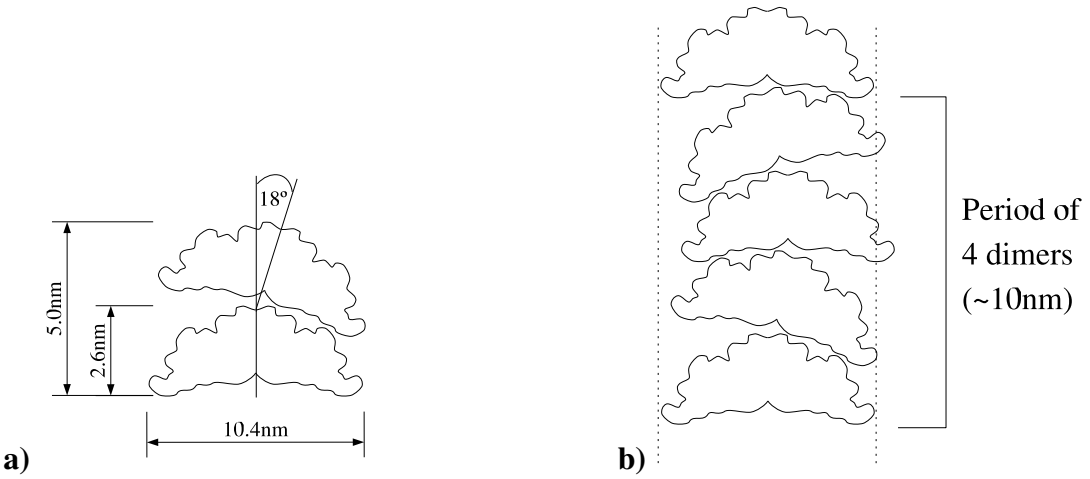


Figure 7.6: Scale diagram of ocr dimers showing a potential top-to-bottom stacking alignment for fibril formation. **a)** To gain the maximum surface area contact an angle of around 18° is required between dimers (for a simple flat tape model). **b)** A minimum period of 4 dimers is required to remain in a straight line.

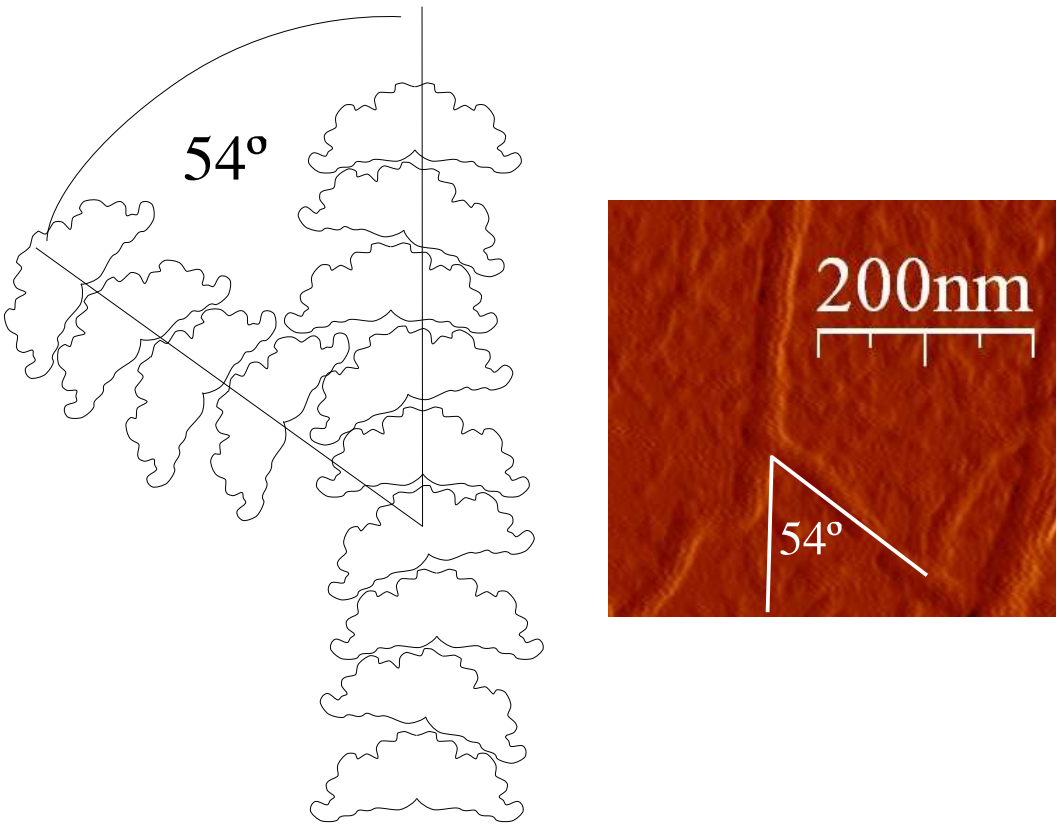
in the alignment of dimers in the fibril or mechanical forces bending the fibrils when samples dry on the AFM substrate surface.

If the tilt of a number of consecutive dimers fall in the same direction a bend develops in the fibril. AFM images show fibrils curving in the x-y plane and regular variations in the height indicate bends along the z-axis (where the plane of figure 7.6 is now z-y rather than x-y). In particular, the repetitions observed in figure 5.26 can be described as a fibril oscillating perpendicularly to the plane. A scale diagram of ocr dimers in figure 7.9 shows that this can be achieved with a period of 12 proteins, reproducing the measured repetition distance of 30nm. The bend is given by 6 consecutive tilts in the same direction. This has a rather low probability (1/16), which may explain why it is not more commonly observed on ocr fibrils.

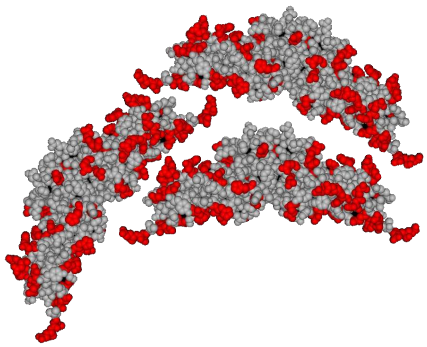
7.4.2 Alternative fibril structures

There are, naturally, a number of other possible configurations for ocr dimers in a fibril, all of which allow a potential binding alignment between acidic residues. It may even be that a variety of different fibril configurations co-exist in a single gel. Two more, alternative fibril structures are considered here.

A variation of the top-to-bottom stacking described above is shown in figure 7.10.



a)



b)

Figure 7.7: **a)** Scale diagram of ocr proteins forming a fibril. The alignment of dimers allows a potential branch to develop at 54. This matches the angle observed in an AFM image of an ocr gel (figure 5.27). **b)** Diagram of the junction showing the acidic residues (red). These align for same-side and opposite-side monomers, allowing both types of binding between proteins. The ammonium sulphate clusters bridge the gaps between acidic residues.

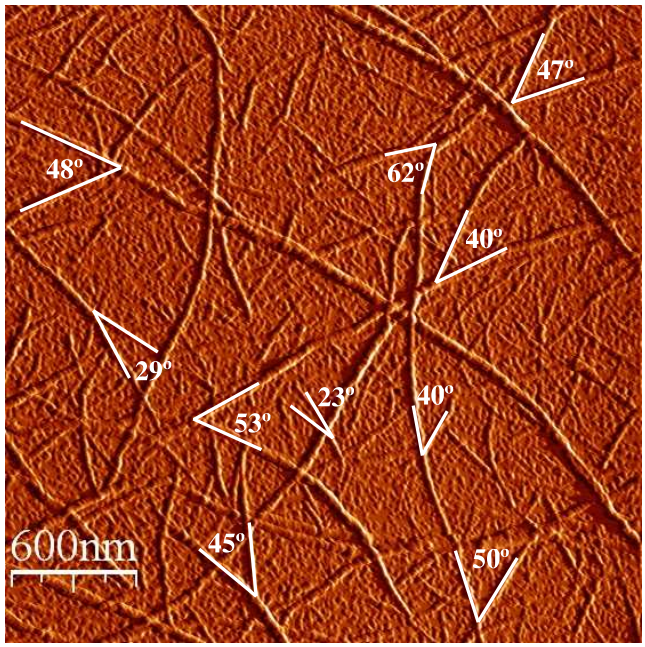


Figure 7.8: A number of possible junctions have been highlighted on this AFM image of a 3.2M ammonium sulphate gel (figure 5.21). They range in angle from around 20° - 60°.

Instead of tilting at 18° there is now an angle of 40° between dimers. This allows a branch to develop using the same binding orientation as in the fibril core, in contrast to figure 7.7. However, the staggered binding means the fibril cross-section dimensions are around 2.6nm×15.6nm (1.5 times the ocr dimer length). This is significantly larger than the maximum fibril height of 10.4nm observed using AFM (figure 5.23).

If, instead of aligning in the same direction, the ocr dimers alternate their orientation they may align along their length, as shown in figure 7.11. This can be achieved in two ways, either through binding between their bottoms (a) or tops (b). Both have possible alignments between acidic residues and allow the fibril to bend but neither the branching nor the observed dimensions can be achieved. The cross-section dimensions for this arrangement are around 2.6nm×5nm, smaller than the fibrils seen in AFM.

Although there are other, more complex, arrangements that exist, the most convincing, and simplest, model for the structure of ocr fibrils appears to involve top-to-bottom stacking, along the bow axis of the dimers. This is capable of producing the features observed in AFM images of ocr gels.

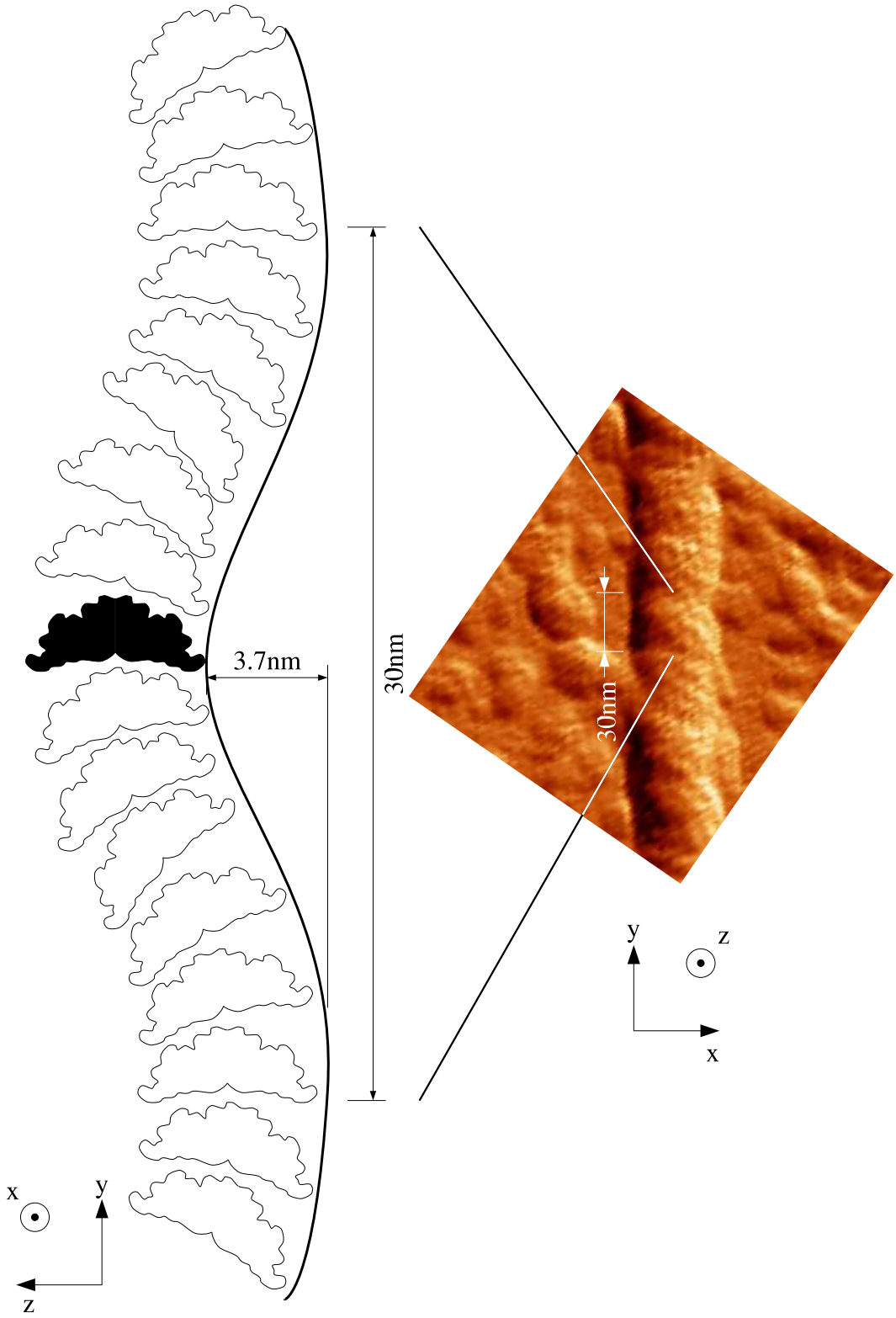


Figure 7.9: A scale diagram of ocr dimers forming an oscillating fibril that matches height oscillations observed in AFM images. 12 dimers achieve a period of 30nm with 6 consecutive dimers tilting in the same direction. The diagram on the left represents a side-on view of the fibril in the AFM image, where the black dimer is the peak of the vertical oscillations.

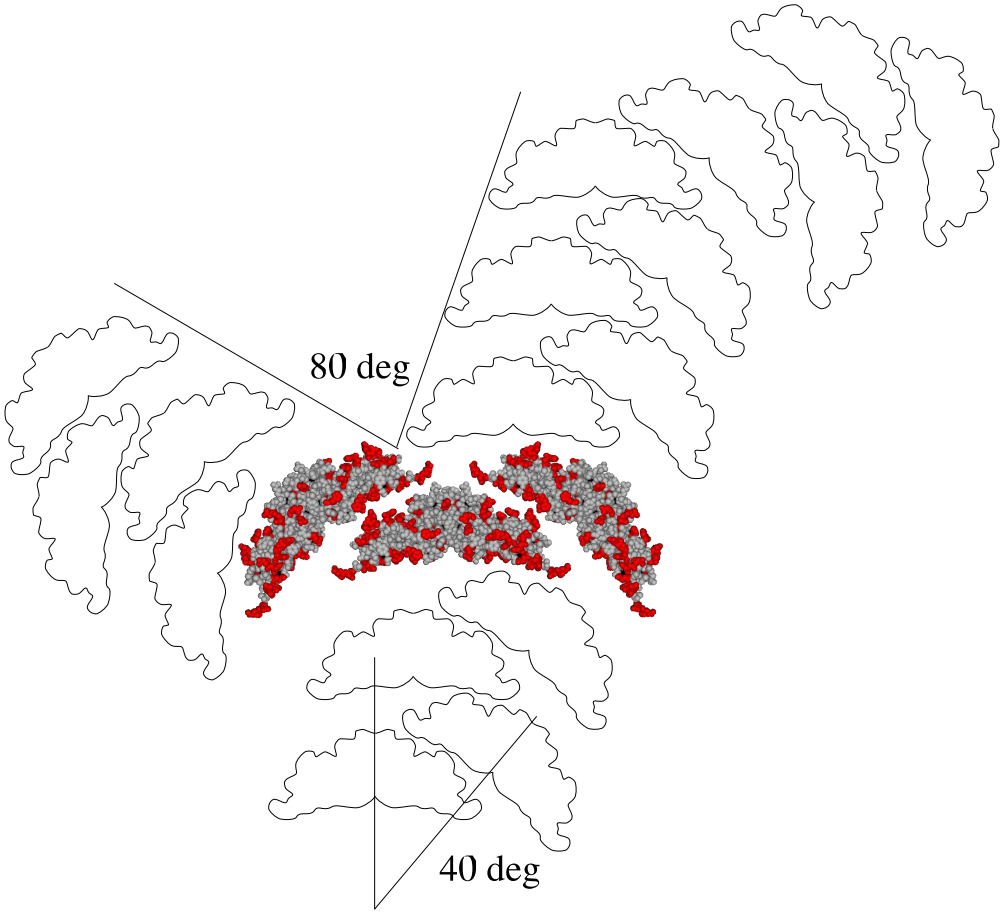


Figure 7.10: A scale diagram of ocr dimers forming a fibril by stacking on either side of the previous dimer. This is similar to the top-to-bottom stacking described in figure 7.6 but with an angle of around 40° rather than 18°. Branches can now be formed using the same binding orientation as in the fibril core but the dimensions of the fibril are larger than that observed with AFM. The three centre dimers show how the acidic residues may align to form bonds.

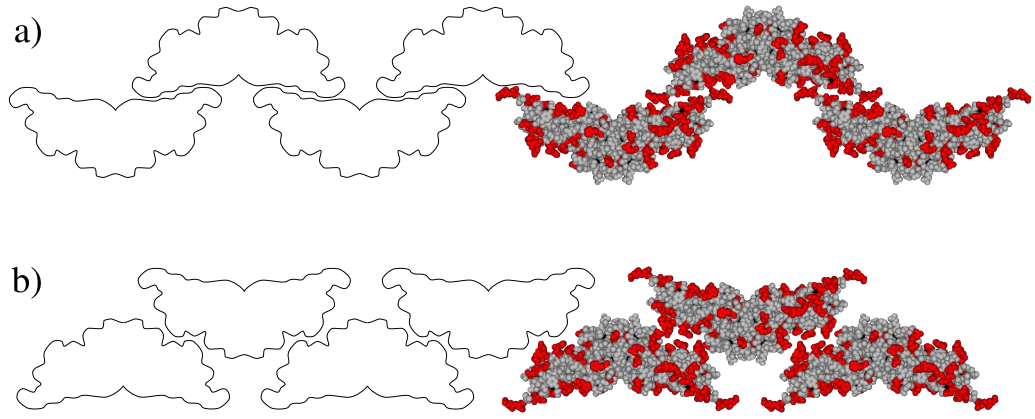


Figure 7.11: A scale diagram of ocr dimers forming a fibril by aligning head to toe. This may be achieved through binding between the bottoms or tops of the dimers. The dimers on the right how the acidic residues may align to form bonds. This does not produce the dimensions of the fibril observed with AFM and it is unclear how branching would occur if binding is limited to this method.

7.5 Other protein fibrils and gels

There are many other protein systems that form similar structures to those discussed above. In this section we compare a number of protein fibril and gel systems, and in particular contrast their circular dichroism and rheological behaviour.

7.5.1 Circular Dichroism and β-sheets

An important question for studying the formation of an ocr gel is whether the protein denatures at such high salt concentrations. The far UV CD spectra for ocr in ammonium sulphate at concentrations up to 3.3M are very similar. Convincing evidence that ocr remains in its native state even within the gel phase. Above 3.4M ammonium sulphate the magnitude of the molecular ellipticity decreases by up to a factor of 3, indicating a slight loss in structure. However, the dichroism remains significantly larger than the fully denatured spectrum and their characteristic shape remains. Of more importance than the intensity of a far UV CD signal are the locations of the minima and maxima. In particular, the minima observed in the spectra at 208nm and 222nm remain, showing that the predominantly α-helical structure is unperturbed up to 3.3M ammonium sulphate. No minima were discovered at 215nm in any of the spectra,

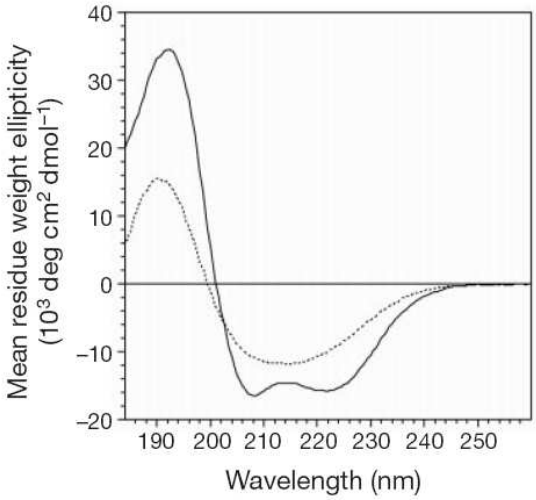


Figure 7.12: Far UV CD spectra for native myoglobin at 22°C (solid line) and fibrils formed after incubating a solution of myoglobin in 50mM sodium borate, pH 9.0 at 65°C for 25 days (dotted line). The strong α-helix signals at 208nm and 222nm have been replaced by a single minima at 215nm, characteristic of β-sheets.

where they would be expected for β-sheets, nor at 198nm for random coils. Figure 7.12 shows an example of the protein myoglobin changing its secondary structure due to a change in solvent conditions [30]. Myoglobin is a highly soluble, compact protein, normally associated with muscles. It has no β-sheet motifs in its secondary structure, which consists mainly of well-defined α-helices. Indeed, far UV CD of the native protein show two distinct minima at 208nm and 222nm. However, after incubating a solution of myoglobin in 50mM sodium borate, pH 9.0 at 65°C for 25 days a large number of fibrils were observed. These gave rise to a single minima at 215nm in the far UV CD spectrum, revealing a predominantly β-sheet structure. The fibrils were later discovered to have a cross-β structure indistinguishable from the pathogenic amyloid fibrils.

CD has been used in many other systems to track the development of β-sheets as proteins form fibrils. The prion protein and β-amyloid protein are both instrumental in neurodegenerative diseases like Creutzfeldt-Jakob disease (CJD) and Alzheimer's disease as they form insoluble fibrils in the subject's brain. Terzi *et al.* studied peptide fragment 25-35 of the β-amyloid protein [143], which forms an antiparallel β-sheet structure. The CD revealed random coil to β-sheet transitions at pH 4.0 (figure 7.13a) and pH 5.5 depending on the protein concentration. At pH 7.4 the solution was dom-

inated by β-sheet fibrils regardless of the concentration. They suggest this is due to the methionine-35 carboxyl group having its highest charge at pH 7.4 and increasing the stability of the β-sheet structure. At high peptide concentrations the characteristic CD minimum of β-sheet structures appears at 215nm. This minimum is shifted to nearer 200nm at lower concentrations, where a random coil structure is predominantly formed. There are no significant absorptions at 208nm or 222nm, indicating an absence of α-helices. At pH 7.0 the spectra for each sample resembles the high concentration traces (3,4 and 5) shown in figure 7.13a).

A similar result was found by Nguyen *et al.* who isolated two peptide fragments (H1: 109-122 and 104H1: 104-122) that were known to be α-helical regions in the prion protein [144]. In an aqueous buffer with 30% acetonitrile, H1 converted to β-sheets and aggregated as amyloid fibrils while the longer peptide 104H1 adopted a random coil conformation and did not form fibrils. The hydrophobic end of H1 is thought to stabilize the β-sheet conformation of the fibrils but this effect is reduced in 104H1 by including extra hydrophobic residues. The characteristic CD spectra for random coil and β-sheets are clearly evident in the spectra obtained for H1 and 104H1 (figure 7.13b).

Ocr does not contain any β-sheet structure, neither in its native nor gelled state, yet it still forms long fibrils under high salt concentrations. These preliminary observations can be supported by calculating the secondary structure ratios. However, caution must be used as there are a number of limitations that exist for far UV CD analysis. α-helices have a strong CD signal and a distinctive spectrum. Their fractional composition can be more accurately determined than that of β-sheets, which have weak CD signals that depend on the lengths of the β-sheets and their spatial orientation [113]. Although far UV CD predominantly involves peptide bonds, other chromophores can also absorb in this region. In particular, aromatic rings may contribute significantly to the spectrum. This effect cannot be accounted for accurately as it is sensitive to the molecular environment of the chromophores. When analysing the data using the online programs it is common practice to analyse the spectra with as many models and reference sets as possible to get the most accurate picture of the secondary structure. Unfortunately we were limited to a single model, and hence only one reference set, as the data below 200nm was particularly noisy, due to the ammonium sulphate. All other models are based on data measured down to 190nm, which provides twice as

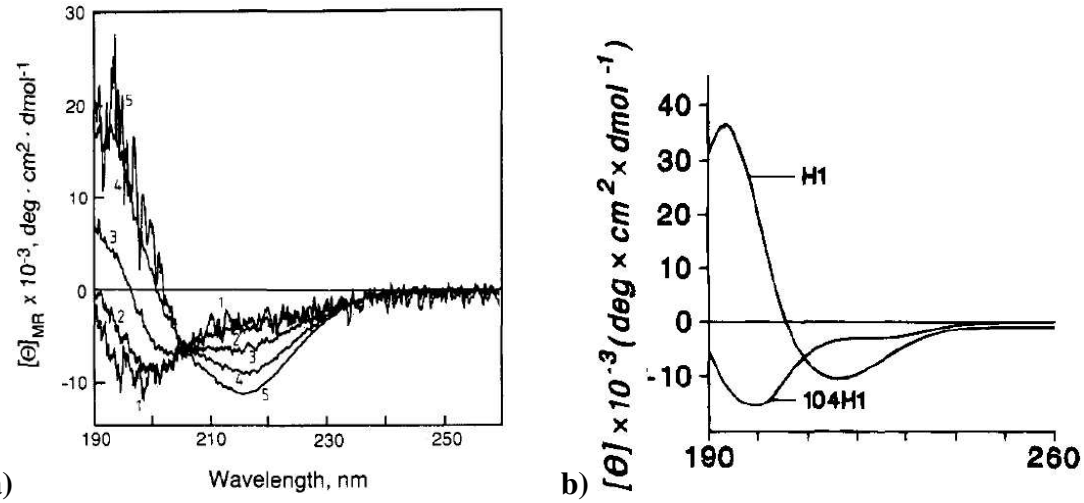
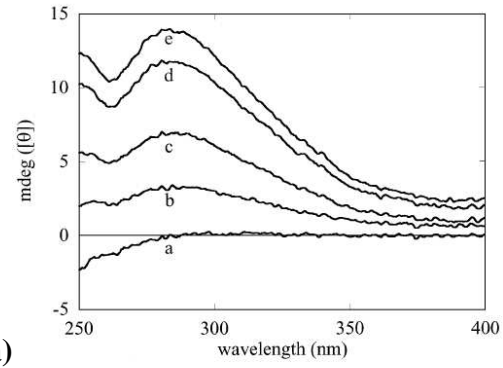


Figure 7.13: **a)** Circular dichroism spectra of peptide fragment 25-35 of the β-amyloid protein in 5mM sodium acetate buffer at pH 4.0. The spectra show a protein concentration dependent transition from random coil (1: 31μM, 2: 62.5μM) to β-sheet (3: 125μM, 4: 250μM, 5: 500μM) [144]. **b)** Two peptide fragments from the prion protein. H1 showing a classic β-sheet motif and 104H1, which has an extra 5 residues, showing a classic random coil motif. Only H1 aggregated into amyloid fibrils [144].

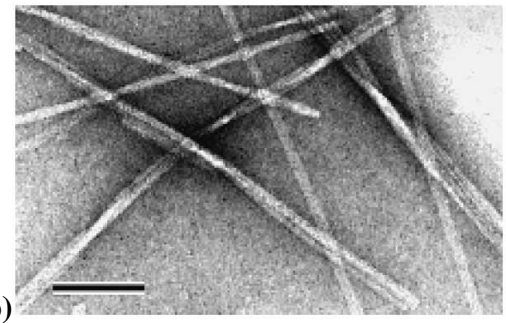
much information for analysis [112]. The models implement a constraint where the fractions of all structure elements must sum to 1.0. A drawback of this for limited or poor quality data is that it is difficult to distinguish a good analysis from a bad one. Particularly as other minor influences, like beta-turns, are often bundled together with "random coil".

Although the secondary structure ratios calculated for ocr are not 100% accurate (for example, a β-sheet ratio of 0.1 is incorrect), it is clear that ammonium sulphate has very little effect on the protein conformation up to 3.3M. The formation of fibrils, therefore, cannot be due to a cross-β structure as in the case of amyloid and other protein fibrils. From 3.4M ammonium sulphate the increased salt concentration causes some of the α-helices to unfold into random coils. This is accompanied by a visual weakening of the gel and a deterioration of the fibrils.

A far more striking effect is observed at near UV wavelengths, where there is a sudden inversion and increase in CD signal at 3.2M ammonium sulphate. Free ammonium sulphate has no molecular ellipticity so the increase in ellipticity could indicate that chirality has been induced by the process of salt binding to ocr [110]. Alter-



a)



b)

Figure 7.14: **a)** Near UV CD spectra of the peptide fragment NAC(1-13), showing an increase in left-handed chirality as fibrils are formed over the period of a week. 1mM NAC at pH 2.2, (a), was incubated for (b) 1 day, (c) 3 days, (d) 5 days and (e) 7 days at 37°C. **b)** Electron micrograph of fibrils formed from 1mM NAC(1-13) after incubating for 1 week. The scale bar represents 100nm. Figures reproduced from [146].

natively the salt may displace water from clefts in the tertiary structure causing the aromatic residues to reorientate slightly. Glazer and Simmons showed that adding N-acetylglucosamine (NAG), a competitive inhibitor, to lysozyme doubles the near UV CD peaks at 285 and 295nm [145]. This wavelength region is characteristic of tryptophan side chains, whereas the peak between 272 and 282nm in figure 5.14a) is in the tyrosine absorption region. Disulphide bonds between cysteine residues can also account for a large proportion of the CD spectra but there are no cysteine molecules in ocr. A similar case was discovered by Abe and Nakanishi [146] using a non-amyloid β-component peptide fragment. It has no aromatic residues or cysteine amino acids yet displays a CD peak at 285nm after a period of incubation. Immediately after the sample was prepared there was no CD signal. The sample was incubated in 1mM NAC, pH 2.2, at 37°C for one week and over this time the peak in molecular ellipticity gradually increased (figure 7.14). They attribute this appearance of chirality to the amyloid fibrils that form over the course of the week. A similar effect may be the cause of the large CD change at 3.2M ammonium sulphate and suggests ocr forms a left-handed chiral structure in the gel phase.

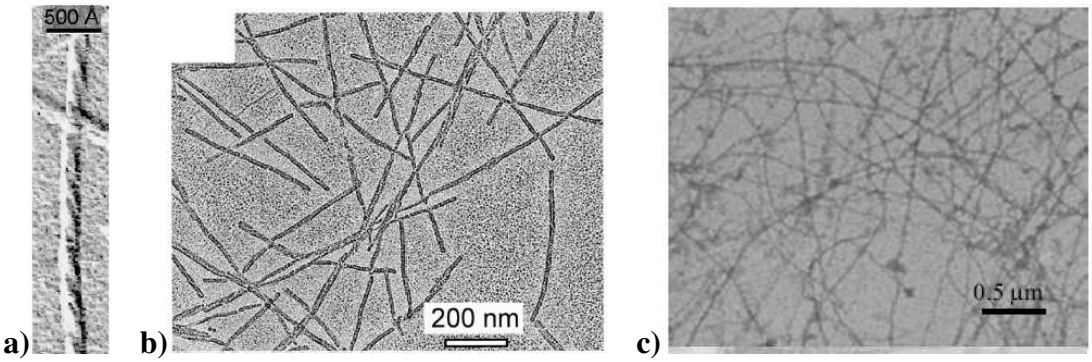


Figure 7.15: Examples of fibril forming systems: **a)** Insulin fibril showing left-handed twist of protofilaments. Formed by heating 2mM insulin at pH 2 to 60°C for 2 days [25]. **b)** P_{II}-I peptides (containing mainly glutamine residues) adopt β-sheet structures and form helical fibrils at a concentration of 6.2mM [24]. **c)** Fibrils formed by heating 2%(w/w) β-lactoglobulin at pH 2 to 80°C for 8h [148].

7.5.2 Twisted tapes

Amyloid fibrils also adopt a left-handed helical structure. This is thought to be due to the left-handed nature of amino acids in proteins that form fibrils [147]. The left-handedness of insulin fibrils has been observed using AFM and CD. In its native state the polypeptide hormone insulin is an α-helical structure with two polypeptide chains linked by disulphide bonds. They form fibrils at high insulin concentrations (2mM) and low pH (around 2) when heated at 60°C for over 2 days, then incubated at room temperature for a few more days. Jimenez *et al.* found that once the α-helices have completely unfolded, the sequence adopts a cross-β structure, which allows proteins to aggregate into protofilaments around 3-4nm in diameter [25]. These coil round each other in a helical manner to form fibrils up to 20nm in diameter with a faintly observable periodicity. Where the twist could be discerned it was consistently found to be left-handed (figure 7.15a).

Dzwolak *et al.* observed the binding of thioflavin T to insulin fibres using circular dichroism in the near UV region [147]. Thioflavin T is an optically-inactive fluorophore that has a high affinity to amyloid fibrils. When bound, they revealed a left-handed, chiral assembly following that of the fibrils. The induced optical activity, which increased with thioflavin T content, may result from altering the conformation of single molecules to produce chirality but is more likely from assembling the achiral

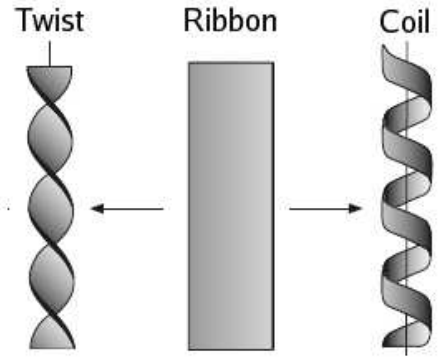


Figure 7.16: Two possible helical conformations that may be adopted from a ribbon. The twisted helix minimises free energy with respect to size [149].

molecules into helical structures.

The vast majority of peptides and proteins that form fibrils do so via a β-sheet motif. Once the peptides have adopted a β-sheet conformation they self-assemble into ribbons, which can twist or coil into left-handed helical structures (figure 7.16) [149, 150]. The former conformation is most likely in fluid materials as it minimises the bulk free energy with respect to size. These structures soften twist round each other to form fibrils of well-defined widths (figure 7.17 a and b). The process is concentration dependent and often takes hours or days to produce significant quantities of fibrils. In experiments with peptide P_{II}-I (consisting of mainly glutamine residues with an arginine and glutamate to prevent random peptide association) ribbon formation was observed at a peptide concentration of 0.2mM and at 6.2mM helical fibrils 8nm wide were seen (figure 7.15b) [24].

The intrinsic twist helps to stabilise the fibrils and prevent them from precipitating. Fibrils (in particular amyloids) commonly have a finite, almost monodisperse thickness of $\sim 7-12$nm. This limited aggregation behaviour resembles that of amphiphilic micellization rather than the infinite aggregation usually expected for attractions between homogeneous molecules. Nyrkova *et al.* suggest this is the result of chiral structures being formed [151]. The helical β-sheet tapes entwine until the energy gain from face-to-face attractions between ribbons is balanced by the energy cost of distorting the tapes to form a twisted fibril.

Twisted fibrils are also able to form stable dispersions due to their helical nature. Stiff, rod-like molecules would be expected to aggregate and precipitate from the solution. When two cylinders come into contact their area of interaction is far greater

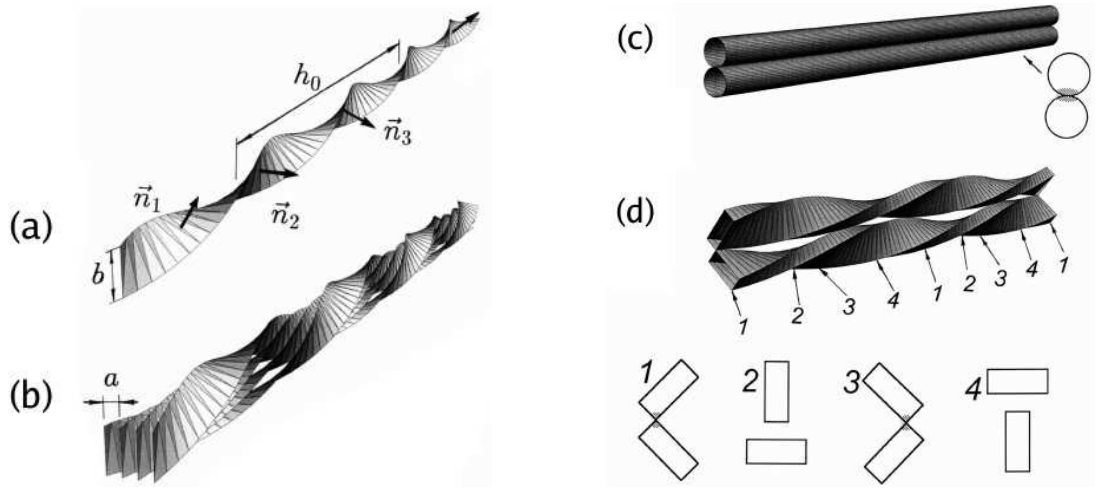


Figure 7.17: (a) Helical tape formed from a stack of peptides or proteins. Ribbon thickness b, width b, helix step $h_0$and surface normal $\vec{n}$. (b) Four ribbons entwined in a stack, where their axis have been slightly distorted to pack together. Note that the ribbons are shown as infinitely thin sheets for clarity. (c) Parallel cylinders can interact along the whole of their length while fibrils (d) only interact occasionally (cross-sections 1 and 3) along their length, so are less prone to aggregation. Diagram reproduced from [151].

than that of two helices (figure 7.17 c and d). This results in a much stronger attraction between rod-like polymers than between twisted fibrils and so the rods are more likely to aggregate. This may explain why bundles of fibrils are seen in AFM images of ocr in 3.1M, 3.4M and 3.5M ammonium sulphate (5.22). At these salt concentrations the CD spectra show little or no helicity. The fibrils formed may not be twisted and may therefore be more prone to aggregate into bundles.

7.5.3 Gel formation

The most common method of forming gels from proteins such as whey, lysozyme and soy is by heating them (typically to around 80°C). The proteins aggregate through patches exposed by their thermal unfolding, and consequently a gel forms as the solution cools down. This is in contrast to the ocr gel, which has been seen to form at temperatures as low as 4°C. An alternative cold gelation process has been investigated for β-lactoglobulin (the main constituent of whey protein) [148, 152, 153]. A heating step is still required but a low concentration of protein is used so they denature but do

not gel upon cooling. A gel is induced at low temperatures by adding salt. This can produce more transparent and more elastic gels with better water-holding capacities than heat-induced ones.

The properties of the β-lactoglobulin gel are largely governed by the choice of salt, pH and ionic strength. Salt ions screen electrostatic interactions between charged proteins. At high pH and ionic strength β-lactoglobulin forms a turbid, particulate gel as random aggregation is favoured under these conditions. If a low pH is used long, linear fibrils with a β-sheet structure are formed (figure 7.15c). They are then cross-linked using salt to form a transparent gel. Ca^{2+} appears to be the most effective cation as it forms intermolecular salt bridges between the negatively charged carboxyl groups of β-lactoglobulin, producing an ordered network of fibrils. Heating a 2%(w/w) solution of β-lactoglobulin at pH 2 to 80°C for 8h, letting it cool then adding 0.01-0.08M $CaCl_2$ resulted in fibrils 2-7μm long [148].

Long, semi-flexible β-sheet tapes have been seen to produce transparent, viscoelastic gels in a number of systems. Peptide molecules resembling segments of native proteins like IsK adopt an antiparallel β-sheet motif in amphiphilic solvents like methanol and 2-chloroethanol [154]. Even at low concentrations (~1mg/ml, 0.2% volume fraction) they aggregate into distinct fibril-like tapes that become highly entangled and form a gel. The tapes are around 8nm wide, corresponding to the length of the peptide β-sheet used) and can be over 0.1μm in length. The peptide fibrils are structurally similar to other protein fibrils including amyloids and the polyglutamine-rich Huntington's protein. The peptide gels formed are similar to biopolymer gels like actin, amylose and gelatin.

The circular dichroism for this system reveals that the peptide secondary structure is predominantly α-helical in 30% methanol but becomes β-sheet in 100% methanol where a gel forms (figure 7.18b). Attractive forces between hydrophobic side-chains help to stabilise the β-sheets and gel structure. Rheology performed on the gel shows that G' is approximately constant over the range of small-strain oscillatory shear applied (figure 7.18a). This viscoelastic behaviour is similar to ocr and other highly entangled polymer gels.

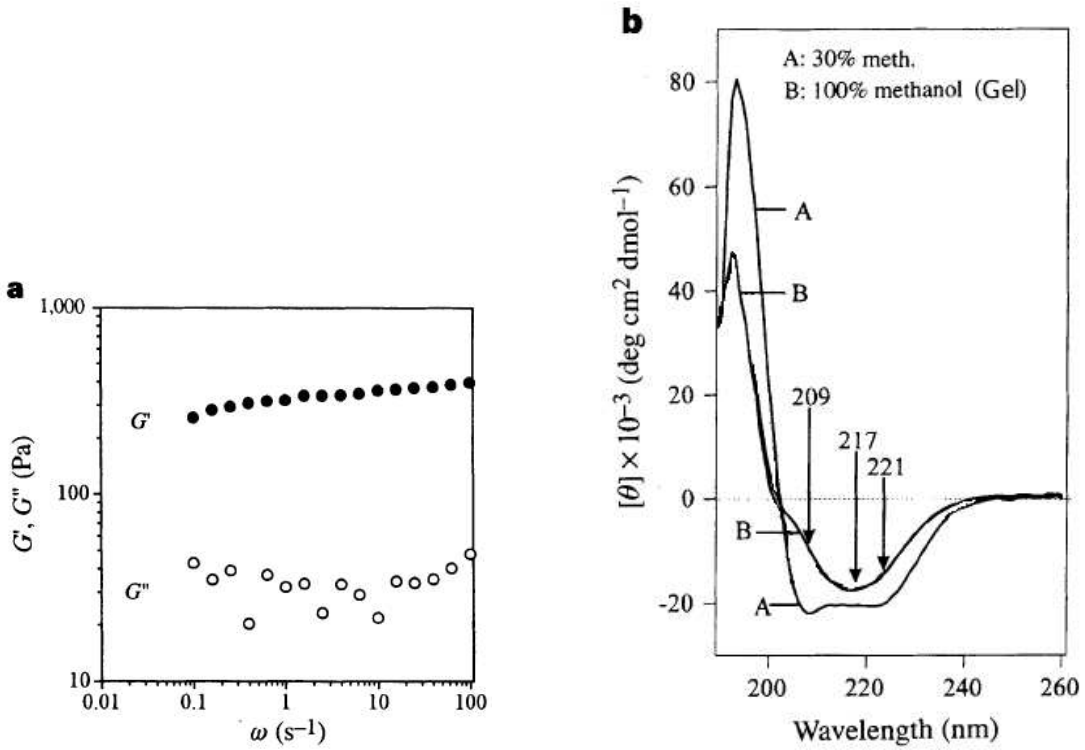


Figure 7.18: **a)** Oscillatory rheology of peptide K24 (a fragment of the IsK protein) at 0.019 volume fraction in 2-chloroethanol at 25°C. The elastic modulus G' is consistently higher than the viscous modulus G", as expected for a viscoelastic gel. **b)** Far UV CD of peptide K24 in 30% and 100% methanol at peptide volume fraction ~0.006%. A gel forms at 100% methanol where the secondary structure has changed from α-helices to β-sheets. Diagrams reproduced from [154].

7.5.4 Actin filaments

Perhaps one of the most similar systems to ocr fibrils are actin filaments. In particular, both are formed upon the addition of salt cations and do not rely on a β-sheet motif to aggregate. Actin polymerization is initiated by neutralisation of negative charges on the actin monomer. The cations interact electrostatically with several, specific sites on the monomer linking them in a chain. Different salts have varying affinities (Mg^{2+} > Ca^{2+} > K^{+}) but actin polymerization is not limited to a specific salt, unlike ocr. The choice of salt, valency and solution pH affect the self-assembly rate of actin filaments but not their length [155]. The higher the salt concentration and greater the affinity, the faster the reaction. Monovalent salts require two orders of magnitude higher concentration than divalent ones (10mM as opposed to 0.1mM) [42]. The most significant increase in rate, however, is obtained by lowering the pH.

Wang et al. measured the fluorescence signal of actin and salt solutions over time [155]. As filaments grew the fluorescence intensity increased. Using this they calculated the diffusion constant and filament length as a function of time. Actin concentrations of the order 10μM were used. Filament formation was seen to occur at room temperature over a time scale of a few minutes at low pH (~6) to a few hours at high pH (~8). The pI of actin is ~5.5 so by decreasing the pH the protein becomes less negative and less cations are required to screen it. The monomers interact more readily in this case and so the polymerization is accelerated. The reduction in protein charge due to a decrease in pH alters the interaction between ocr molecules in a similar manner but leads to the proteins precipitating.

Even under the most efficient conditions actin polymerisation is significantly slower than ocr fibril formation. Like amyloid fibrils actin filaments require a nucleation stage. Actin dimers are thermodynamically unstable but above a critical number of monomer aggregates the addition of subsequent units becomes more probable than dissociation.

Although a larger protein concentration is required, along with a substantially lower salt concentration, actin filaments themselves are comparable in size to ocr. They are around 5-10nm in diameter, relatively straight and are commonly greater than 1μm long. The monomers align in two staggered, parallel rows that are twisted into a right-handed helix (figure 7.19) [132]. Its periodicity is 37nm, which is similar to that seen in ocr fibrils, although the chirality is opposite. Filaments formed from 1mg/ml actin are seen to be evenly distributed [43, 40]. They overlap or entwine with other filaments

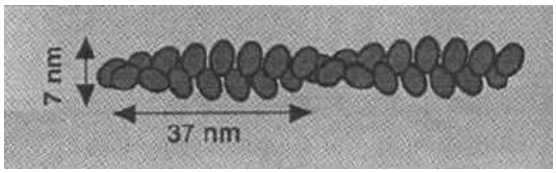


Figure 7.19: Diagram of the right-handed helix formed from two staggered rows of actin monomers [132]. The filament is around 5-10nm wide, has a periodicity of 37nm and has lengths of over 1μm long. The conformation is similar to that envisaged for ocr fibrils where a single chain of monomers, linked by ammonium sulphate molecules, forms a left-handed helix.

with an average separation of 0.7μm and at random angles. These junction points are responsible for the formation of a weak actin gel even though there are no specific chemical cross-links. Unfortunately there was no information available about whether the filamentous gel is transparent or opaque.

When actin forms filaments in the presence of actin-binding proteins, stronger gels are produced. Their underlying network may be isotropic or anisotropic. The protein abp [1] exists in numerous cell types and promotes cross-linking at right-angles. This results in a mesh of short, straight filaments separated by ~0.1μm [43]. On the other hand, the protein scruin binds to long filaments, allows them to bunch together and forms cross-links between bundles of multiple filaments [41, 42]. Both types of gel have a high elasticity that can vary by several decades, 0.1-300 Pa, over a wide frequency range, 0.005-30 rad/s, depending on the level of cross-linking. Like ocr gels, rheology of actin gels show that G' is consistently greater than G" and has little frequency dependence (figure 7.20).

However, for all its similarities to the ocr fibrillar gel, actin is able to form fibrils with a variety of salts and does not rely on a specific cation to link protein molecules. Indeed, all the fibrillar systems reviewed above have significant differences to ocr, despite many forming left-handed fibrils and transparent gels. The most important difference is the ability of ocr to resist aggregation at high salt concentrations and form fibrils without changing its secondary structure.

[1] actin binding protein

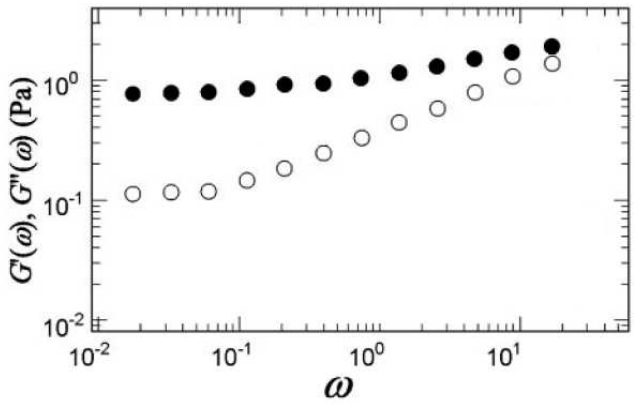


Figure 7.20: Rheology of a 11.9μM filamentous actin solution prepared with 0.36μM scruin (an actin binding protein), pH 7.5. G' (●) is dominant over G" (○) for a wide range of frequencies, indicating an elastic gel network [41].

Chapter 8

Conclusion

Interest in ocr began in microbiology labs. By mimicing DNA, ocr is a powerful anti-restriction protein, and the only such protein for which the structure is known. However, a routine check during purification revealed unexpected behaviour and threw ocr into the fields of physics and biochemistry. Ocr resists precipitation in concentrated salt solutions and instead forms a transparent gel when ammonium sulphate is added. Both the important biological role and physical behaviour of ocr stem from its specific physical properties: the banana-like shape and high negative charge.

In this thesis, we set out to investigate the mechanism behind this unusual behaviour. Protein aggregation can occur reversibly or irreversibly. The latter commonly involves partially unfolding the protein by submitting it to extreme conditions (raising the temperature, for example) and can lead to well-defined structures forming, like amyloid fibrils. The reversible process may be more gentle and involve no unfolding but it produces only amorphous aggregates. Ocr is unusual as it forms a gel through a reversible process while achieving a characteristic fibrillar structure.

Protein precipitation is typically achieved by increasing the salt concentration of a protein solution. The salt ions screen the surface charges, exposing attractive patches and allowing the proteins to approach close enough to aggregate. Ocr is able to resist this process because of its large negative charge, -56e, most of which is distributed over the surface in a helical motif. No precipitation is observed in solutions up to 3.2M ammonium sulphate (80% saturation), whereupon a transparent gel rapidly forms. Above 3.4M, and towards saturation, the gel appears to weaken, indicating a loss of gel structure and possible protein unfolding.

Standard, globular proteins like lysozyme typically precipitate in solutions with

ionic strengths of the order 0.6M. Ocr would require ionic strengths 50 times greater, at physiological pH 8. Since this is unachievably high, the only way to precipitate ocr is by reducing its charge. At pH 5.6, ocr has lost 15% of its charge and begins to aggregate. By pH 4 it has almost completely precipitated.

Ionic strength, however, is not the only controlling factor. While no salt is able to precipitate or denature ocr (at pH 8), only ammonium sulphate causes it to form a gel. These two important observations are clearly revealed with circular dichroism. At far UV wavelengths, no other salt produces a change in the CD spectra from that of native ocr. The spectra for increasing ammonium sulphate concentrations are also very similar up to 3.4M, where they decrease slightly in magnitude. Secondary structure analysis of the data indicates a slight unfolding of α-helices into random coils above 3.4M ammonium sulphate, which may explain the observed weakening of the gel. Near UV CD of ocr is also unaffected by the choice of salt, except at 3.2M and 3.3M ammonium sulphate where the spectra suddenly invert and increase by an order of magnitude. The ocr proteins have adopted a left-handed helical structure. This must have at least one dimension smaller than optical wavelengths to remain transparent, and must be spatially extended in another dimension to form a gel at low protein concentrations.

Such properties are exhibited by fibril systems and, indeed, ocr fibrils are observed in AFM images of ocr gels. A network of well distributed fibrils is clearly seen spanning the gel. They ranged in height from 2.6nm to 10.4nm (the thickness and width of an ocr protein) and were commonly over 1μm long. At salt concentrations just outside the gel phase, poorly-formed fibrils were observed, along with amorphous ocr aggregates.

Proteins typically form fibrils by unfolding slightly, either to reveal an existing β-sheet structure or to change their native structure into β-sheets, through which they self-assemble into long, rigid fibrils. Since no β-sheets are observed in the analysis of the CD spectra, ocr must form fibrils by another means. Only ammonium sulphate is able to gel ocr, implying a unique interaction is required to link the proteins. We propose that ammonium sulphate forms a specific salt bridge between adjacent ocr proteins via electrostatic interactions between the ammonia of the salt and the carboxyl groups on the surface of the proteins. A high salt concentration keeps the salt ions associated amidst decreasing entropy. The most likely orientation of proteins in the

fibril is a top-to-bottom stacking, along the bow axis of ocr. This gives the observed fibril dimensions and allows them to bend and branch, reproducing the behaviour of ocr fibrils observed in AFM images.

Future work

Naturally, any research project is restricted in its scope by the availability of apparatus, materials and time. This has been especially true with ocr. The lengthly period required to produce manageable quantities of purified protein, coupled with resource-hungry experiments, necessitates the performance of only a selected subset of experiments.

There are many more conditions under which the behaviour of ocr may be investigated. The mechanical properties of the gel may be further characterised by rheology, varying salt and ocr concentrations at different oscillatory stresses and pH values. The gel has been seen to form between 4-25°C but has only been investigated at room temperature. Its viscoelastic properties under increasing temperature would be interesting to study. More extensive AFM should also be performed under these varying conditions. In particular, a wider range of pH values and ocr concentrations, and a number of different salts to confirm fibrils only form with ammonium sulphate.

An important aspect of the ocr gel, which this thesis has only touched on, is its reversibility. This is certainly an area that needs further investigation. An insightful experiment would be to take the UV absorbance and CD spectra of a 3.2M ammonium sulphate ocr gel, remove the salt either by dialysis or dilution, repeat the spectroscopy measurements then reform the gel by dialysis and take the measurements a third time. This should clearly reveal the extent of the reversibility of ocr fibrillation and gelation. The transition may also be exhibited in other proteins, in particular those with DNA mimicking properties and surface carboxylic acids.

It would be interesting to perform x-ray and small-angle light scattering to probe the length scales of the gel and fibrils, which could be compared to AFM results and possible fibril models. Raman spectroscopy may provide a more accurate measure of the secondary structure content in various salt concentrations. It would be useful to confirm that ocr is not denaturing, thereby limiting fibril models to those involving only native ocr dimers as subunits. Transition Electron Microscopy (TEM) would give a mass per unit length measurement for the fibrils, which may help identify the way subunits are connected. TEM could also be used to measure the dimensions of fibrils

in a “wet” environment, avoiding concerns about removing water and collapsing fibrils by drying samples. Gluteraldahyde would be added to a gel to fix the protein structures and allow the salted buffer to be removed without affecting the fibrils. AFM using a wet cell would achieve this too, if a large enough sample was prepared and the fibrils naturally adheared to the surface. Force measurements could also be performed with an AFM. Pulling the tip through fibrils of different dimensions to calculate the force required to break them may help narrow the possible configurations of proteins in the fibril. There are undoubtedly many other types of experiments available for the study of ocr. This thesis represents the first step in characterising the novel formation of fibrils from the unique protein and will hopefully motivate further investigations in this field.

Appendix A

Protocols for SDS-PAGE

SDS polyacrylamide gel electrophoresis is a technique for separating proteins based on their molecular weight. It is explained in section 3.1.1. This appendix will describe the methods (A.1) and materials (A.2) required to perform such an experiment.

A.1 Method

A.1.1 Prepare gel

The gel is prepared between two glass plates separated with two 0.75mm spacers, one on each side, and locked into a casting stand. Sealing gel is pipetted down beside both spacers on the inside of the glass sandwich, until it forms a layer around 5mm thick at the bottom. Once this has set the separating gel is added up to around 20mm below the top of the glass plates. Add some water to the top of the separating gel as it encourages a horizontal edge to form and the gel sets very slowly when exposed to air. The water can gently be poured off once the separating gel has set. Add the stacking gel to the top of the plates then imediately insert a 15mm tall teflon comb in between the glass sandwich. The teeth on the comb will create the wells in which the protein samples are loaded and normally come with 10 or 15 teeth.

A.1.2 Run samples

Add an equal volume of $2\times$SDS loading buffer to each of the samples. Place the plates and casting stand into the base tray and fill it with $1\times$SDS/electrophoresis buffer.

Slowly and carefully remove comb and rinse the wells with buffer. Load the samples into the well with a pipette, ensuring no samples overflow into neighbouring wells. Depending on the protein concentration of the samples, use around 4μl per well for a 15-teeth-comb (10μl max) and around 10μl per well for a 10-teeth-comb (30μl max). Cover with the lid and run for 25-30mA per gel until the smallest proteins (identifed by the bromophenol blue marker in the loading buffer) reached the bottom of the gel (30-60 minutes).

A.1.3 Stain and destain

Once the electrophoresis has finished carefully prise the glass plates appart and slide the gel into a tray containing coomassie blue stain solution. Glently shake until the protein bands are clear (1-2 hours). Transfer the gel to a tray containing destain to remove excess stain. Leave gently shaking until the background gel is transparent (generally over 2 hours, perhaps overnight).

A.2 Materials

A.2.1 Stock solutions and buffers required

- *4× Tris.Cl/SDS, pH 8.8*
- *4× Tris.Cl/SDS, pH 6.8*
- *30% Acrylamide (29:1 Acrylamide:Bisacrylamide)*
- *Temed*
- *10% APS (Ammonium persulfate) [i.e. 10g APS / 100ml H_2O]*
- 2× SDS/loading buffer
- 5× SDS/electrophoresis buffer
- Stain and Destain

A.2.2 Protocols for buffers and stains

4× Tris.Cl/SDS, *pH 6.8*

- to 40ml H_2O add:	
6.05g Tris base	*(0.5M Tris.Cl)*
0.4g SDS	*(0.4%)*
- adjust to **pH 6.8**	
- add H_2O to 100ml	
- filter (0.4μl) and store (4°C)	

4× Tris.Cl/SDS, *pH 8.8*

- to 300ml H_2O add:	
91g Tris base	*(1.5M Tris.Cl)*
2g SDS	*(0.4%)*
- adjust to **pH 8.8**	
- add H_2O to 500ml	
- filter (0.4μl) and store (4°C)	

5× SDS/electrophoresis buffer

15.1g	Tris base
72g	Glycine
5g	SDS
	H_2O to 1L

Do NOT adjust pH (8.3 when diluted).

Dilute to 1× before use.

2× SDS/loading buffer

25ml	4× Tris.Cl/SDS, **pH 6.8**
20ml	Glycerol
4g	SDS
2 ml	2-β-Mercaptoethanol
1mg	Bromphenol blue
	H_2O to 100ml

Stain

0.02g/100ml	Coomassie brilliant blue R-250
30%	Methanol
10%	Acetic acid

Destain

30%	Methanol
10%	Acetic acid

A.2.3 Protocols for gels

Sealing gel

1.65ml	Acrylamide (30%)
1.25ml	4×Tris.Cl/SDS, **pH 8.8**
2.10ml	H_2O
40μl	TEMED
50μl	APS (10%)

Separating gel (15%)

7.50ml	Acrylamide (30%)
3.75ml	4×Tris.Cl/SDS, **pH 8.8**
3.75ml	H_2O
50μl	TEMED
50μl	APS (10%)

Stacking gel

0.65ml	Acrylamide (30%)
1.25ml	4×Tris.Cl/SDS, **pH 6.8**
3.00ml	H_2O
40μl	TEMED
50μl	APS (10%)

A.2.4 Other Separating Gel Compositions

Use 15% gel for OCR (13.5 kDa)

Final Acrylamide conc.	10%	11%	12%	13%	14%	**15%**
Acrylamide (30%)	5.00ml	5.50ml	6.00ml	6.50ml	7.00ml	**7.50ml**
Tris.Cl/SDS, pH 8.8	3.75ml	3.75ml	3.75ml	3.75ml	3.75ml	**3.75ml**
H_2O	6.25ml	5.75ml	5.25ml	4.75ml	4.25ml	**3.75ml**
TEMED	50μl	50μl	50μl	50μl	50μl	**50μl**
APS (10%)	50μl	50μl	50μl	50μl	50μl	**50μl**

List of Figures

List of Tables

Bibliography

[1] M.R. Tock and D.T.F. Dryden. The biology of restriction and anti-restriction. *Current Opinion in Microbiology*, 8:466–472, 2005.

[2] T.E. Hugli. Protease inhibitors: novel and therapeutic application and development. *Trends in biotechnology*, 14:409–412, 1996.

[3] G.M. Whitesides and B. Grzybowski. Self-assembly at all scales. *Science*, 295:2418–2421, 2002.

[4] S. Zhang. Fabrication of novel biomaterials through molecular self-assembly. *Nature Biotechnology*, 21(10):1171–1178, 2003.

[5] P. Mesquida, C.K. Riener, C.E. MacPhee, and R.A. McKendry. Morphology and mechanical stability of amyloid-like peptide fibrils. *Journal of Materials Science: Materials in Medicine*, OnlineFirst, 2007.

[6] C.E. MacPhee and C.M. Dobson. Formation of mixed fibrils demonstrates the generic nature and potential utility of amyloid nanostructures. *J.Am.Chem.Soc.*, 122:12707–12713, 2000.

[7] A. Keller. Aspects of polymer gels. *Faraday Discussions*, 101:1–49, 1995.

[8] H. Rehage and H. Hoffmann. Viscoelastic surfactant solutions; model for rheological research. *Molecular Physics*, 74(5):933–973, 1991.

[9] K. Hanabusa, K. Okui, K. Karaki, M. Kimura, and H. Shirai. Organogels formed by N-Benzyloxycarbonyl-L-alanine 4-Hexadecanoyl-2-nitrophenyl Ester and related compounds. *Journal of Colloid and Interface Science*, 195:86–93, 1997.

[10] R.G. Larson. The structure and rheology of complex fluids. *Oxford University Press, New York*, 1999.

[11] G.M. Kavanagh and S.B. Ross-Murphy. Rheological characterisation of polymer gels. *Prog.Polym.Sci.*, 23:533–562, 1998.

[12] P.H. Hermans and H.R. Kruyt (editor). Colloid Science, Volume II. *Elsevier Publishing Company Inc.*, pages 483–651, 1949.

[13] H. Sedgwick, S.U. Egelhaaf, and W.C.K. Poon. Clusters and gels in systems of sticky particles. *J. Phys.: Condens. Matter*, 16:4913–4922, 2004.

[14] A.H. Clark. Application of cascade theory to the description of microphase-separated biopolymer gels. *Faraday Discussions*, 101:77–91, 1995.

[15] V. Trappe and P. Sandkühler. Colloidal gels - low-density disordered solid-like states. *Current Opinion in Colloid and Interface Science*, 8:494–500, 2004.

[16] T. van Vliet and P. Walstra. Large deformation and fracture behaviour of gels. *Faraday Discussions*, 101:359–370, 1995.

[17] D. Durand, J.C. Gimel, and T. Nicolai. Aggregation, gelation, and phase separation of heat denatured globular proteins. *Physica A*, 304:253–265, 2002.

[18] K.G. de Kruif, M.A.M. Hoffmann, M.E. van Marle, P.J.J.M. van Mil, S.P.F.M. Roefs, M. Verheul, and N. Zoon. Gelation of proteins from milk. *Faraday Discussions*, 101:185–200, 1995.

[19] W.C.K. Poon, A.D. Pirie, and P.N. Pusey. Gelation in colloid-polymer mixtures. *Faraday Discussions*, 101:65–76, 1995.

[20] D.R. Booth, M. Sunde, V. Bellotti, C.V. Robinson, W.L. Hutchinson, P.E. Fraser, P.N. Hawkins, C.M. Dobson, S.E. Radford, C.C.F. Blake, and M.B. Pepys. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature*, 385:787–793, 1997.

[21] D.J. Selkoe. Folding proteins in fatal ways. *Nature*, 426:900–904, 2003.

[22] Wetzel. Ideas of order for amyloid fibril structure. *Structure*, 10:1031–1036, 2002.

[23] C. Goldsbury, J. Kistler, U. Aebi, T. Arvinte, and G.J.S. Cooper. Watching amyloid fibrils grow by time-lapse atomic force microscopy. *J. Mol. Biol.*, 285:33–39, 1999.

[24] A. Aggeli, I.A. Nyrkova, M. Bell, R. Harding, L. Carrick, T.C.B. McLeish, A.N. Semenov, and N. Boden. Hierarchical self-assembly of chiral rod-like molecules as a model for peptide β-sheet tapes, ribbons, fibrils and fibers. *Proc. Natl. Acad. Sci.*, 98(21):11857–11862, 2001.

[25] J.L. Jiménez, E.J. Nettleton, M. Bouchard, C.V. Robinson, C.M. Dobson, and H.R. Saibil. The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci.*, 99(14):9196–9201, 2002.

[26] J.I. Guijarro, M. Sunde, J.A. Jones, I.D. Campbell, and C.M. Dobson. Amyloid fibril formation by an SH3 domain. *Biochemistry*, 95:4224–4228, 1998.

[27] J.T. Jarrett and P.T.Jr Lansbury. Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry*, 31:12345–12352, 1992.

[28] J. Zurdo, J.I. Guijarro, J.L. Jimenez, H.R. Saibil, and C.M. Dobson. Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. *Journal of Molecular Biology*, 311:325–340, 2001.

[29] C.M. Dobson. Protein folding and misfolding. *Nature*, 426:884–890, 2003.

[30] M. Fändrich, M.A. Fletcher, and C.M. Dobson. Amyloid fibrils from muscle myoglobin. *Nature*, 410:165–166, 2001.

[31] E. Scherzinger, A. Sittler, K. Schweiger, V. Heiser, R. Lurz, R. Hasenbank, G.P. Bates, H. Lehrach, and E.E. Wanker. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc. Natl. Acad. Sci.*, 96:4604–4609, 1999.

[32] S. Chen, V. Berthelier, J.B. Hamilton, B. O'Nuallain, and R. Wetzel. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry*, 41:2002, 7391-7399.

[33] C.A. Ross, M.A. Poirier, E.E. Wanker, and M. Amzel. Polyglutamine fibrillogenesis: the pathway unfolds. *Proc. Natl. Acad. Sci. Commentary*, 100(1):1–3, 2003.

[34] A.D. Parkhe, S.K. Seeley, K. Gardner, L. Thompson, and R.V. Lewis. Structural studies of spider silk proteins in the fiber. *Journal of Molecular Recognition*, 10:1–6, 1997.

[35] S. Putthanarat, N. Stribeck, S.A. Fossey, R.K. Eby, and W.W. Adams. Investigation of the nanofibrils of silk fibers. *Polymer*, 41:7735–7747, 2000.

[36] Y. Shen, A. Johnson, and D.C. Martin. Microstructural characterization of bombyx mori silk fibers. *Macromolecules*, 31, 1998.

[37] J. Zhang, Y.-L. Wang, L. Gu, and J. Pan. Atomic force microscopy of actin. *Acta Biochimica et Biophysica Sinica*, 35(6):489–494, 2003.

[38] T. Lehto, M. Miaczynska, M. Zerial, D.J. Müller, and F. Severin. Observing the growth of individual actin filaments in cell extracts by time-lapse atomic force microscopy. *FEBS Letters*, 551:25–28, 2003.

[39] E.A. Brotschi, J.H. Hartwig, and T.P. Stossel. The gelation of actin by actin-binding protein. *Journal of Biological Chemistry*, 253(24):8988–8993, 1978.

[40] Z. Shao, D. Shi, and A.V. Somlyo. Cryoatomic force microscopy of filamentous actin. *Biophysical Journal*, 78:950–958, 2000.

[41] J.H. Shin, M.L. Gardel, L. Mahadevan, P. Matsudaira, and D.A. Weitz. Relating microstructure to rheology of a bundled and cross-linked f-actin network in vitro. *Proc. Natl. Acad. Sci.*, 101(26):9636–9641, 2004.

[42] T.P. Stossel, C. Chaponnier, R.M. Ezzell, J.H. Hartwig, P.A. Janmey, D.J. Kwiatkowski, S.E. Lind, D.B. Smith, F.S. Southwick, H.L. Yin, and K.S. Zaner. Nonmuscle actin-binding proteins. *Ann. Rev. Cell Biol.*, 1:353–402, 1985.

[43] R. Niederman, P.C. Amrein, and J. Hartwig. Three-dimensional structure of actin filaments and of an actin gel made with actin-binding protein. *Journal of Cell Biology*, 96:1400–1413, 1983.

[44] R.M.J. Cotterill. Biophysics: an introduction. *John Wiley and Sons Ltd.*, 2002.

[45] D.R. Link, G.N. Natale, R. Shao, J.E. Maclennan, N.A. Clark, E. Körblova, and D.M. Walba. Spontaneous formation of macroscopic chiral domains in a fluid smectic phase of achiral molecules. *Science*, 278:1924–1927, 1997.

[46] J. Thisayukta, H. Takezoe, and J. Watanabe. Study on helical structure of the B_4 phase formed from achiral banana-shaped molecule. *Jpn. J. Appl. Phys.*, 40:3277–3287, 2001.

[47] T. Niori, T. Sekine, J. Watanabe, T. Furukawa, and H. Takezoe. Distinct ferroelectric smectic liquid crystals consisting of banana shaped achiral molecules. *J. Mater. Chem.*, 6(7):1231–1233, 1996.

[48] J. Barberá, N. Gimeno, I. Pintre, M.B. Ros, and J.L. Serrano. Self-assembled bent-core side-chain liquid crystalline polymers. *Chem. Commun.*, 11:1212–1214, 2006.

[49] G. Heppke and D. Moro. Liquid crystals: chiral order from achiral molecules. *Science*, 279(5358):1872–1873, 1998.

[50] D. Shen, S. Diele, I. Wirt, and C. Tschierske. A novel class of non-chiral banana-shaped liquid crystals with ferroelectric properties. *Chem. Commun.*, pages 2573–2574, 1998.

[51] J.W. Goodby, D.W. Bruce, M. Hird, C. Imrie, and M. Neal. An introduction to materials discussion No.4: molecular topology in liquid crystals. *J. Mater. Chem.*, 11:2631–2636, 2001.

[52] J. Thisayukta, Y. Nakayama, S. Kawauchi, H. Takezoe, and J. Watanabe. Distinct formation of a chiral smectic phase in achiral banana-shaped molecules with a central core based on a 2,7-dihydroxynaphthalene unit. *J. Am. Chem. Soc.*, 122:7441–7448, 2000.

[53] R.A. Copeland. Methods for protein analysis. *Chapman and Hall*, 1994.

[54] C. Branden and J. Tooze. Introduction to protein structure (2nd ed.). *Garland Publishing*, 1999.

[55] C. Atanasiu, O. Byron, H. McMiken, S.S. Sturrock, and D.T.F. Dryden. Characterisation of the structure of OCR, the gene 0.3 protein of bacteriophage T7. *Nucleic Acids Res.*, 29(14):3059–3068, 2001.

[56] D.T.F. Dryden. Reeling in the bases. *Nature Structural & Molecular Biology*, 11(9):804–806, 2004.

[57] N.E. Murray. Type I restriction systems: sophisticated molecular machines. *Microbio. Mol. Biol. Rev.*, 64(2):412–434, 2000.

[58] N.E. Murray. Immigration control of DNA in bacteria: self versus non-self. *Microbiology*, 148:3–20, 2002.

[59] T.A. Bickle and D.H. Krüger. Biology of DNA restriction. *Microbiological Reviews*, 57(2):434–450, 1993.

[60] K.K. Mark and F.W. Studier. Purification of the gene 0.3 protein of bacteriophage T7, an inhibitor of the DNA restriction system of Escherichia coli. *J. Biol. Chem.*, 256(5):2573–2578, 1981.

[61] S.S. Sturrock, D.T.F. Dryden, C. Atanasiu, J. Dornan, S. Bruce, A. Cronshaw, P. Taylor, and M.D. Walkinshaw. Crystallization and preliminary x-ray analysis of OCR, the product of gene 0.3 of bacteriophage T7. *Acta Cryst. D*, 57(11):1652–1654, 2001.

[62] C. Atanasiu, T.J. Su, S.S. Sturrock, and D.T.F. Dryden. Interaction of the OCR gene 0.3 protein of bacteriophage T7 with EcoKI restriction/modification enzyme. *Nucleic Acids Res.*, 30(18):3936–3944, 2002.

[63] P.K. Bandyopadhyay, F.W. Studier, D.L. Hamilton, and R. Yuan. Inhibition of the type I restriction-modification enzymes EcoB and EcoK by the gene 0.3 protein of bacteriophage T7. *J. Mol. Biol.*, 182:567–578, 1985.

[64] J.J. Blackstock, S.U. Egelhaaf, C. Atanasiu, D.T.F. Dryden, and W.C.K. Poon. Shape of OCR, the gene 0.3 protein of bacteriophage T7: modeling based on light scattering experiments. *Biochemistry*, 40(33):9944–9949, 2001.

[65] J.J. Dunn, M. Elzinga, K.K. Mark, and F.W. Studier. Amino acid sequence of the gene 0.3 protein of bacteriophage T7 and nucleotide sequence of its mRNA. *J. Biol. Chem.*, 256(5):2579–2585, 1981.

[66] M.D. Walkinshaw, P. Taylor, S.S. Sturrock, C. Atanasiu, T. Berge, R.M. Henderson, J.M. Edwardson, and D.T.F. Dryden. Structure of OCR from bacteriophage T7, a protein that mimics B-form DNA. *Mol. Cell.*, 9(1):187–194, 2002.

[67] P.W. Atkins. Physical chemistry. *Oxford University Press*, 6th ed., 1998.

[68] H.N. Po and N.M. Senozan. The Henderson-Hasselbalch equation: its history and limitations. *Journal of Chemical Education*, 78(11):1499–1503, 2001.

[69] C. Mohan. Buffers. *Calbiochem*, EMD Biosciences Inc, 2003.

[70] K.A. Dill and S. Bromberg. Molecular driving forces. *Garland Science*, 2003.

[71] J.N. Israelachvili. Intermolecular and surface forces. *Academic Press*, 1985.

[72] D.H. Everett. Basic principles of colloid science. *Royal Society of Chemistry*, 1994.

[73] Helen Sedgwick. Colloidal metastability. *PhD Thesis*, The University of Edinburgh, 2003.

[74] T.E. Creighton. Proteins: structures and molecular principles. *W.H. Freeman and Co.*, 2nd ed., 1993.

[75] J.L. Abernethy. Franz Hofmeister: the impact of his life and research on chemistry. *Proceedings of the California Association of Chemistry Teachers*, 44(3):177–180, 1967.

[76] M. Boström, B. Lonetti, E. Fratini, P. Baglioni, and B.W. Ninham. Why pH titration in protein solutions follows a Hofmeister series. *J. Phys. Chem. B*, 110:7563–7566, 2006.

[77] D.E. Kuehner, J. Engmann, F. Fergg, M. Wernick, H.W. Blanch, and J.M. Prausnitz. Lysozyme net charge and ion binding in concentratedaqueous electrolyte solutions. *Journal of Physical Chemistry B*, 103:1368–1374, 1999.

[78] R. Chang. Physical chemistry with applications to biological systems. *Macmillan Publishing Co.Inc.*, New York, 1981.

[79] C. Tanford and R. Roxby. Interpretation of protein titration curves. Application to lysozyme. *Biochemistry*, 11(11):2192–2198, 1972.

[80] R.E. Canfield. The amino acid sequence of egg white lysozyme. *Journal of Biological Chemistry*, 238(8):2698–2707, 1963.

[81] D.C. Phillips. The hen egg-white lysozyme molecule. *N.A.S. Symposium*, 57:484–495, 1967.

[82] M. Ullner and B. Jönsson. A monte carlo study of titrating polyelectrolytes in the presence of salt. *Macromolecules*, 29:6645–6655, 1996.

[83] C. Labbez, B. Jönsson, I. Pochard, A. Nonat, and B. Cabane. Surface charge density and electrokinetic potential of highly charged minerals: experiments and monte carlo simulations on calcium silicate hydrate. *J. Phys. Chem. B*, 110:9219–9230, 2006.

[84] R.P. Sear. The effects of added salt on the second virial coefficients of the complete proteome of *e.coli*. *Journal of Chemical Physics*, 118(11):5157, 2003.

[85] R.F. Weaver and P.W. Hedric. Basic genetics. *Wm.C.Brown Publishers*, 2nd Edition, 1995.

[86] R.K. Scopes. Protein purification. *Springer-Verlag*, 3rd Edition, 1994.

[87] C. Atanasiu. Characterization of OCR, the product of gene 0.3 from bacteriophage T7. *PhD Thesis*, The University of Edinburgh, 2000.

[88] S. Erramilli, F. Österberg, and B. Vogelaar. Undergraduate laboratory: principles of gel electrophoresis. *Am.J.Phys.*, 63(7):639–642, 1995.

[89] C. Viney and R.A. Fenton. Physics and gel electrophoresis: using terminal velocity to characterize molecular weight. *Eur.J.Phys.*, 19:575–580, 1998.

[90] M.P. (editor) Deutscher. Guide to protein purification. *Methods in Enzymology*, 182:894, 1990.

[91] U.K. Laemmli. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227:680–685, 1970.

[92] D.B. Wetlaufer. Ultraviolet spectra of proteins and amino acids. *Adv. in Prot. Chem.*, 17:303–390, 1962.

[93] A. Cooper. Biophysical chemistry. *Royal Society of Chemistry*, 2004.

[94] N.C. Price. Enzymology labfax. *Academic Press, Inc. and BIOS Scientific Publishers Ltd.*, 1996.

[95] O Warburg and W Christian. Isolation and crystallization of enolase. *Biochem.Z*, 1942.

[96] pET System. Gene expression protocols. *Novagen*, 7th Edition, 1997.

[97] G. Karp. Cell biology. *McGraw-Hill Inc.*, 2nd ed., 1984.

[98] B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. Essential cell biology. *Garland Publishing Inc.*, 1998.

[99] H.A. Barnes, J.F. Hutton, and K. Walters. An introduction to rheology. *Elsevier Science Publishers*, 1993.

[100] F.A. Morrison. Understanding rheology. *Oxford University Press*, 2001.

[101] M. Reiner. *Physics Today*, 17(62), 1969.

[102] M.E. Cates. Structural relaxation and rheology of soft condensed matter. *Slow Dynamics and Nonequilibrium Relaxations in Condensed Matter, editors J-L.Barrat, M.Feigelman, J.Kurchan, and J.Dalibar*, EDP Sciences-Springer:74–130, 2003.

[103] J.F. Steffe. Rheological methods in food process engineering. *Freeman Press*, 2nd edition, 1996.

[104] V.T. O'Brien and M.E. Mackay. Stress components and shear thickening of concentrated hard sphere suspensions. *Langmuir*, 16:7931–7938, 2000.

[105] P.H.T. Uhlherr, J. Guo, C. Tiu, X.-M. Zhang, J.Z.-Q Zhou, and T.-N. Fang. The shear-induced solid-liquid transition in yield stress materials with chemically different structures. *J. Non-Newtonian Fluid Mech.*, 125:101–119, 2005.

[106] A.K. Sood, R. Bandyopadhyay, and G. Basappa. Linear and nonlinear rheology of wormlike micelles. *Pramana Journal of Physics*, 53(1):233–235, 1999.

[107] H.A. Barnes and D. Bell. Controlled-stress rotational rheometry: an historical review. *Korea-Australia Rheology Journal*, 15(4):187–196, 2003.

[108] T. McLeish. Rheology of linear and branched polymers. *Soft and Fragile Matter, editors M.E.Cates and M.R.Evans*, Institute of Physics Publishing:79–111, 2000.

[109] A. Rodger and B. Norden. Circular dichroism and linear dichroism. *Oxford University Press*, 1997.

[110] S.M. Kelly and N.C. Price. The use of circular dichroism in the investigation of protein structure and function. *Current Protein and Peptide Science*, 1:349–384, 2000.

[111] S.M. Kelly and N.C. Price. The application of circular dichroism to studies of protein folding and unfolding. *Biochim Biophys Acta.*, 1338(2):161–185, 1997.

[112] W.C. Johnson. Secondary structure of proteins through circular dichroism spectroscopy. *Annual review of biophysics and biomolecular structure*, 17:145–166, 1988.

[113] T.E. Creighton (editor). Protein structure a practical approach. *IRL Press*, 1990.

[114] L. Whitmore and B.A. Wallace. Dichroweb, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.*, 32:W668–W673, 2004.

[115] A. Lobley, L. Whitmore, and B.A. Wallace. Dichroweb: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, 18:211–212, 2002.

[116] M.A. Andrade, P. Chacon, J.J. Merelo, and F. Moran. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng.*, 6:282–290, 1993.

[117] N.J. Greenfield. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem.*, 235:1–10, 1996.

[118] G. Binnig, C.F. Quate, and Ch. Gerber. Atomic force microscope. *Physical Review Letters*, 56(9):930–933, 1986.

[119] H.G. Hansma and L. Pietrasanta. Atomic force microscopy and other scanning probe microscopies. *Current Opinion in Chemical Biology*, 2:579–584, 1998.

[120] A.K. Chamberlain, C.E. MacPhee, J. Zurdo, L.A. Morozova-Roche, H.A.O. Hill, C.M. Dobson, and J.J. Davis. Ultrastructural organization of amyloid fibrils by atomic force microscopy. *Biophysical Journal*, 79:3282–3293, 2000.

[121] D.J. Ellis, D.T.F. Dryden, T. Berge, J.M. Edwardson, and R.M. Henderson. Direct observation of DNA translocation and cleavage by the *ecoki* endonuclease using atomic force microscopy. *Nature Structural biology*, 6(1):15–17, 1999.

[122] H.G. Hansma and J.H. Hoh. Biomolecular imaging with the atomic force microscope. *Annu. Rev. Biophys. Biomol. Struct.*, 23:115–39, 1994.

[123] L. Haggerty and A.M. Lenhoff. STM and AFM in biotechnology. *Biotechnol. Prog.*, 9:1–11, 1993.

[124] Cora-Ann Schoenenberger Andreas Engel and Daniel J Müller. High resolution imaging of native biological sample surfaces using scanning probe microscopy. *Current Opinion in Structural Biology*, 7:297–284, 1997.

[125] J. Zhang, C.R. Daubert, and E.A. Foegeding. Characterization of polyacrylamide gels as an elastic model for food gels. *Rheol Acta*, 44:622–630, 2005.

[126] T. Gisler, R.C. Ball, and D.A. Weitz. Strain hardening of fractal colloidal gels. *Phys.Rev.Lett.*, 82(5):1064–1067, February 1999.

[127] S. Ikeda and K. Nishinari. On solid-like rheological behaviour of globular protein solutions. *Food Hydrocolloids*, 15:401–406, 2001.

[128] S.B. Ross-Murphy. Structure-property relationships in food biopolymer gels and solutions. *J.Rheol.*, 39(6):1451–1463, 1995.

[129] A.H. Clark, G.M. Kavanagh, and S.B. Ross-Murphy. Globular protein gelation - theory and experiment. *Food Hydrocolloids*, 15:383–400, 2001.

[130] M.M. Ould Eleya, S. Ko, and S. Gunasekaran. Scaling and fractal analysis of viscoelastic properties of heat-induced protein gels. *Food Hydrocolloids*, 18:313–323, 2004.

[131] P. Pouzot, L. Benyahia, and T. Nicolai. Dynamic mechanical characterization of the heat-induced formation of fractal globular protein gels. *J. Rheol.*, 48(5):1123–1134, 2004.

[132] F.C. MacKintosh, J. Kas, and P.A. Janmey. Elasticity of semiflexible biopolymer networks. *Phys.Rev.Lett.*, 75(24):4425–4428, 1995.

[133] T. Coviello, K. Kajiwara, W. Burchard, M. Dentini, and V. Crescenzi. Solution properties of xanthan. dynamic and static light scattering from native and modified xanthans in dilute solutions. *Macromolecules*, 19:2826–2831, 1986.

[134] B.I. Kankia, V. Buckin, and V.A. Bloomfield. Hexamminecobalt(III)-induced condensation of calf thymus DNA: circular dichroism and hydration measurements. *Nucleic Acids Res.*, 29(13):2795–2801, 2001.

[135] J.R. Lakowicz. Principles of fluorescence spectroscopy. *Plenum Press*, 1983.

[136] G. Qian and M. Wang. Study on the microstructural evolution of silica gel during sol-gel-gel-glass conversions using the fluorescence polarization of rhodamine B. *J.Phys.D:Appl.Phys.*, 32:2462–2466, 1999.

[137] J.C. Butler, T. Angelini, J.X. Tang, and G.C.L. Wong. Ion multivalence and like-charge polyelectrolyte attraction. *Physical Review Letters*, 91(2), 2003.

[138] P.E. Smith. Computer simulation of cosolvent effects on hydrophobic hydration. *Journal of Physical Chemistry B*, 103:525–534, 1999.

[139] W.C.K. Poon, S.U. Egelhaaf, P.A. Beales, A. Salonen, and L. Sawyer. Protein crystallization: scaling of charge and salt concentration in lysozyme solutions.

Journal of Physics: Condensed Matter, Letter to the Editor, 12:L569 – L574, 2000.

[140] H. Sedgwick, K. Kroy, A. Salonen, M.B. Robertson, S.U. Egelhaaf, and W.C.K. Poon. Non-equilibrium behavior of sticky colloidal particles: beads, clusters and gels. *Journal of Physics: Condensed Matter, Letter to the Editor*, 12:L569 – L574, 2000.

[141] A. George and W.W. Wilson. Predicting protein crystallization from a dilute solution property. *Acta Cryst.*, D50:361–365, 1994.

[142] D. Rosenbaum, P.C. Zamora, and C.F. Zukoski. Phase behavior of small attractive colloidal particles. *Physical Review Letters*, 76(1):150–153, 1996.

[143] E. Terzi, G. Hölzemann, and J. Seelig. Reversible random coil-β-sheet transition of the Alzheimer β-amyloid fragment (25-35). *Biochemistry*, 33:1345–1350, 1994.

[144] J. Nguyen, M.A. Baldwin, F.E. Cohen, and S.B. Prusiner. Prion protein peptides induce α-helix to β-sheet conformational transitions. *Biochemistry*, 34:4186–4192, 1995.

[145] A.N. Glazer and N.S. Simmons. Structure-function relationships in lysozyme by circular dichroism. *J. Amer. Chem. Soc.*, 88:2335–2336, 1966.

[146] H. Abe and H. Nakanishi. Novel observation of a circular dichroism band originating from amyloid fibril. *Analytical Sciences*, 19:171–173, 2003.

[147] W. Dzwolak and M. Pecul. Chiral bias of amyloid fibrils revealed by the twisted conformation of Thioflavin T: an induced circular dichroism/DFT study. *FEBS Letters*, 579:6601–6603, 2005.

[148] C. Veerman, H. Ruis, L.M.C. Sagis, and E. van der Linden. Effect of electrostatic interactions on the percolation concentration of fibrillar β-lactoglobulin gels. *Biomacromolecules*, 3:869–873, 2002.

[149] R. Oda, I. Huc, M. Schmutz, S.J. Candau, and F.C. MacKintosh. Tuning bilayer twist using chiral counterions. *Nature*, 399:566–569, 1999.

[150] I.A. Nyrkova, A.N. Semenov, A. Aggeli, M. Bell, N. Boden, and T.C.B. McLeish. Self-assembly and structure transformations in living polymers forming fibrils. *European Physical Journal B*, 17:499–513, 2000.

[151] I.A. Nyrkova, A.N. Semenov, A. Aggeli, and N. Boden. Fibril stability in solutions of twisted β-sheet peptides: a new kind of micellization in chiral systems. *European Physical Journal B*, 17:481–497, 2000.

[152] P. Hongsprabhas, S. Barbut, and A.G. Marangoni. The structure of cold-set whey protein isolate gels prepared with Ca^{2+}. *Lebensm.-Wiss. u.-Technol.*, 32:196–202, 1999.

[153] C. Veerman, H. Baptist, L.M.C. Sagis, and E. van der Linden. A new multistep Ca^{2+}-induced cold gelation process for β-lactoglobulin. *Journal of Agricultural and Food Chemistry*, 51:3880–3885, 2003.

[154] A. Aggeli, M. Bell, N. Boden, J.N. Keen, P.F. Knowles, T.C.B. McLeish, M. Pitkeathly, and S.E. Radford. Responsive gels formed by the spontaneous self-assembly of peptides into polymeric β-sheets tapes. *Nature*, 386:259–262, 1997.

[155] F. Wang, R.V. Sampogna, and B.R. Ware. pH dependence of actin self-assembly. *J. Biophys.*, 55(1989):293–298.